Search Notes

10/677,865

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2006, 22:33:20 ; Search time 8124 Seconds
(without alignments)
11125.195 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1588.4	99.9	1590	6	AR429544	AR429544 Sequence
2	1586.8	99.8	1725	6	AX786044	AX786044 Sequence
3	1586.8	99.8	1725	8	AF030130	AF030130 Homo sapi
4	1582	99.5	1590	6	A84909	A84909 Sequence 2
5	1582	99.5	1590	6	BD072818	BD072818 Polypepti
6	1582	99.5	1725	6	CQ722442	CQ722442 Sequence
7	1579.6	99.3	1590	6	AX233491	AX233491 Sequence
8	1570.8	98.8	2284	6	CS086650	CS086650 Sequence
9	1570	98.7	2530	6	AX336639	AX336639 Sequence
10	1570	98.7	2530	6	AX786041	AX786041 Sequence
11	1570	98.7	2530	8	HUMSPROT	L20815 Human S pro
12	1553.2	97.7	2062	8	BC031993	BC031993 Homo sapi
13	1552.8	97.7	1587	11	BT019904	BT019904 Synthetic
14	1503.2	94.5	1764	8	AF224751	AF224751 Homo sapi
15	1503.2	94.5	1764	8	AF224753	AF224753 Homo sapi
16	1503.2	94.5	2127	8	AF224752	AF224752 Homo sapi
17	1503.2	94.5	4740	6	AX786043	AX786043 Sequence
18	1503.2	94.5	5337	6	BD095296	BD095296 The metho
19	1503.2	94.5	5700	6	AX233490	AX233490 Sequence
20	1503.2	94.5	25235	6	BD095297	BD095297 The metho
c 21	1503.2	94.5	30911	8	AB088114	AB088114 Homo sapi
22	1503.2	94.5	39196	8	AB023060	AB023060 Homo sapi
23	1503.2	94.5	44379	8	AC006163	AC006163 Homo sapi
24	1503.2	94.5	55050	6	AX522117	AX522117 Sequence
c 25	1503.2	94.5	56809	8	CR759805	CR759805 Human DNA
26	1503.2	94.5	110000	8	BA000025_08	Continuation (9 of
c 27	1503.2	94.5	135396	8	BX927139	BX927139 Human DNA
c 28	1503.2	94.5	349980	6	CS039418	CS039418 Sequence
29	1501.6	94.4	1764	8	AF224747	AF224747 Homo sapi
30	1501.6	94.4	2127	8	AF286165	AF286165 Homo sapi
c 31	1501.6	94.4	81418	8	CR753819	CR753819 Human DNA
32	1501.2	94.4	1506	6	AX786098	AX786098 Sequence
33	1500	94.3	1764	8	AF224748	AF224748 Homo sapi
34	1500	94.3	1764	8	AF224749	AF224749 Homo sapi
35	1500	94.3	1764	8	AF224754	AF224754 Homo sapi
c 36	1500	94.3	113388	8	AL773544	AL773544 Human DNA
37	1498.2	94.2	1509	8	AF49132982	AF491330 Homo sapi
38	1488.8	93.6	110000	8	BA000041_03	Continuation (4 of
39	1487.2	93.5	1767	8	AF224757	AF224757 Homo sapi
40	1487.2	93.5	1767	8	AF224758	AF224758 Homo sapi
41	1485.6	93.4	1767	8	AF224756	AF224756 Homo sapi
c 42	1485.6	93.4	204298	8	AL662844	AL662844 Human DNA
43	1484	93.3	1761	8	AF224750	AF224750 Homo sapi
c 44	1484	93.3	3889	8	HSM807103	BX640954 Homo sapi
45	1474.4	92.7	1764	8	AF224755	AF224755 Homo sapi

ALIGNMENTS

RESULT 1
AR429544
LOCUS AR429544 1590 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6645509.
ACCESSION AR429544
VERSION AR429544.1 GI:40189837
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1590)
AUTHORS Serre,G.B.R., Simon,M. and Weber-Vivat,M.
TITLE Polypeptide expressed in the horny layer of epidermis and use thereof
JOURNAL Patent: US 6645509-A 2 11-NOV-2003;
Societe L'Oreal S.A.; Paris;
FRX;
FEATURES Location/Qualifiers
source 1. .1590
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Applicant

Query Match 99.9%; Score 1588.4; DB 6; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
AX786044
LOCUS AX786044 1725 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 14 from Patent WO03050311.
ACCESSION AX786044
VERSION AX786044.1 GI:32953661
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
TITLE Methods and compositions for diagnosing and treating hypotrichosis simplex
JOURNAL Patent: WO 03050311-A 14 19-JUN-2003;
Wella AG (DE)
FEATURES Location/Qualifiers
source 1. .1725
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.8%; Score 1586.8; DB 6; Length 1725;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 3
AF030130
LOCUS AF030130 1725 bp mRNA linear PRI 03-AUG-2000
DEFINITION Homo sapiens corneodesmosin mRNA, complete cds.
ACCESSION AF030130
VERSION AF030130.1 GI:2921269
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1725)
AUTHORS Guerrin,M., Simon,M., Montezin,M., Haftek,M., Vincent,C. and Serre,G.
TITLE Expression cloning of human corneodesmosin proves its identity with the product of the S gene and allows improved characterization of its processing during keratinocyte differentiation
JOURNAL J. Biol. Chem. 273 (35), 22640-22647 (1998)
PUBMED 9712893
REFERENCE 2 (bases 1 to 1725)
AUTHORS Guerrin,M., Simon,M., Montezin,M., Vincent,C. and Serre,G.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) LBPDM, Faculte de Medecine Purpan, 37 Allees Jules Guesde, Toulouse 31000, France
FEATURES Location/Qualifiers
source 1. .1725
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21"
/cell_type="keratinocyte; corneocyte"
/tissue_type="epidermis"
CDS 15. .1604
/note="associated with corneocyte-specific modified desmosomes; late differentiation protein of epidermis; similar to human differentiated keratinocyte S protein encoded by GenBank Accession Number L20815"
/codon_start=1
/product="corneodesmosin"
/protein_id="AAC24196.1"
/db_xref="GI:2921270"
/translation="MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKD
PTRITSPNDPCLTGKGDSSGPSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSS
FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPC
SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSP
GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
PSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGS
KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
LGPQLADPEVFLPQGELLDSP"
ORIGIN

Query Match 99.8%; Score 1586.8; DB 8; Length 1725;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 4
A84909
LOCUS A84909 1590 bp DNA linear PAT 21-JAN-2000
DEFINITION Sequence 2 from Patent WO9844105.
ACCESSION A84909
VERSION A84909.1 GI:6733756
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified sequences.
REFERENCE 1 (bases 1 to 1590)
AUTHORS Simon,M. and Serre,G.B.
TITLE POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS AND USE THEREOF
JOURNAL Patent: WO 9844105-A 2 08-OCT-1998;
OREAL (FR); SIMON MICHEL (FR)
FEATURES Location/Qualifiers
source 1. .1590
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 99.5%; Score 1582; DB 6; Length 1590;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1585; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTATGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

```
RESULT 5
BD072818
LOCUS       BD072818               1590 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Polypeptide expressed in the horny layer of epidermis and use
            thereof.
ACCESSION   BD072818
VERSION     BD072818.1  GI:22618421
KEYWORDS    JP 2001511019-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1590)
  AUTHORS   Serre,G.B.R., Simon,M. and Vivat,M.W.
  TITLE     Polypeptide expressed in the horny layer of epidermis and use
  JOURNAL   Patent: JP 2001511019-A 1 07-AUG-2001;
            L'OREAL
COMMENT     OS   Unidentified
            PN   JP 2001511019-A/1
            PD   07-AUG-2001
            PF   27-MAR-1998 JP 1998541244
            PR   28-MAR-1997 FR     97/03899,11-SEP-1997 FR     97/11317 PI
            GUY BRUNO RENE SERRE,MICHEL SIMON,MARINA WEBER VIVAT PC
            C12N15/12,C07K14/47,A61K38/17,C07K16/18
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Polypeptide expressed in the horny layer of epidermis and use
            CC           thereof
            FH   Key             Location/Qualifiers
            FT   source          1. .1590
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             99.5%;  Score 1582;  DB 6;  Length 1590;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1585;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTATGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy      1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCCAGCCAG 1080

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            |||||||||||||||||||||||||||||||
Db      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 6
CQ722442

```
LOCUS       CQ722442               1725 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8376 from Patent WO02068579.
ACCESSION   CQ722442
VERSION     CQ722442.1  GI:42283299
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 8376 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1725
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             99.5%;  Score 1582;  DB 6;  Length 1725;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1585;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
```

```
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 GGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 7
AX233491
LOCUS AX233491 1590 bp DNA linear PAT 11-SEP-2001
DEFINITION Sequence 134 from Patent WO0162788.
ACCESSION AX233491
VERSION AX233491.1 GI:15592978
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Olaveson,M., Lench,N., Allen,M. and Tazi-Ahnini,R.U.
TITLE Corneodesmosin based test and model for inflammatory disease
JOURNAL Patent: WO 0162788-A 134 30-AUG-2001;
Oxagen Limited (GB)
FEATURES Location/Qualifiers
source 1. .1590
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1. .1590
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC69732.1"
/db_xref="GI:15592979"
/translation="MGSSRAPWMGRVGGHGMXALLLAGLLLPGTLAKSIGTFSDPCKD
PTRITSPNDPCXTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGXSGSHSGSSSSHSSSSSS
FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTXGVSSSGQSVSSNQRPC
SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSP
GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
PSSSRVPSXSSISSSXXXPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPXGS
KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
LGPQLADPEVFLPQGELLBSP"
ORIGIN

Query Match 99.3%; Score 1579.6; DB 6; Length 1590;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1570; Conservative 18; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGWTGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
```

```
Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
            ||:||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db       121 CCYTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCYTCACTGGGAAGGGT 180

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
            ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GACTCCAGCGGYTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TCGGGAARCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||:
Db       541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAR 600

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCYGGCCCCTACATC 720

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGRCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAYCCT 960

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db      1081 TCGGCAGCTTCCTCGGCCATTGCRTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            :||||||||||||||||||||:|||:||:||||||||||||||||||||||||||||||||
Db      1201 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCSAGT 1320

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db      1321 GGCAAAATCATCCTTCAGCCTYGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            |||||||||||||||||:||||||||||||
Db      1561 CCCCAAGGAGAGTTACTCRACAGTCCATAA 1590
```

RESULT 8
CS086650
LOCUS CS086650 2284 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 12 from Patent EP1531161.
ACCESSION CS086650
VERSION CS086650.1 GI:66712119
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Lal,P., Corley,N.C., Guegler,K.J., Patterson,C.,
Baughn,M.R. and Yue,H.
TITLE Human epidermal proteins HEPI-1 to HEPI-6
JOURNAL Patent: EP 1531161-A 12 18-MAY-2005;
Incyte Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. .2284
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 98.8%; Score 1570.8; DB 6; Length 2284;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286
```

```
Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db      1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            |||||||||||||||||||||||||||||||
Db      1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636
```

RESULT 9
AX336639
LOCUS AX336639 2530 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 7148 from Patent WO0194629.
ACCESSION AX336639
VERSION AX336639.1 GI:18127358
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 7148 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1. .2530
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 98.7%; Score 1570; DB 6; Length 2530;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603


RESULT 10
AX786041
LOCUS       AX786041               2530 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequence 11 from Patent WO03050311.
ACCESSION   AX786041
VERSION     AX786041.1  GI:32953659
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
  TITLE     Methods and compositions for diagnosing and treating hypotrichosis
            simplex
  JOURNAL   Patent: WO 03050311-A 11 19-JUN-2003;
            Wella AG (DE)
FEATURES             Location/Qualifiers
     source          1. .2530
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             98.7%;  Score 1570;  DB 6;  Length 2530;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1584;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

```
RESULT 11
HUMSPROT
LOCUS       HUMSPROT                2530 bp    mRNA    linear   PRI 02-SEP-1994
DEFINITION  Human S protein mRNA, complete cds.
ACCESSION   L20815
VERSION     L20815.1  GI:414809
KEYWORDS    S gene; S protein; cell differentiation; keratin; keratinocyte
            differentiation; loricrin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2530)
  AUTHORS   Zhou,Y. and Chaplin,D.D.
  TITLE     Identification in the HLA class I region of a gene expressed late
            in keratinocyte differentiation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 90 (20), 9470-9474 (1993)
   PUBMED   8415725
COMMENT     Original source text: Homo sapiens (library: lambda gt10 (from
            Clontech)) neonate foreskin cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .2530
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="6p21.3"
                     /cell_type="keratinocyte"
                     /tissue_type="foreskin"
                     /dev_stage="neonate"
                     /germline
                     /tissue_lib="lambda gt10 (from Clontech)"
     CDS             63. .1523
                     /standard_name="S protein"
                     /note="epidermal granular cell layer-specific"
                     /codon_start=1
                     /protein_id="AAA21321.1"
                     /db_xref="GI:414810"
                     /translation="MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG
                     DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYS
                     SGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSAL
                     PTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVS
                     HSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKS
                     YGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKY
                     FSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSS
                     GSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLT
                     LTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRIS"
     intron          99^100
                     /number=1
     polyA_signal    2511. .2516
     polyA_site      2530
ORIGIN

  Query Match           98.7%;  Score 1570;  DB 8;  Length 2530;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1584;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

```
RESULT 12
BC031993
LOCUS       BC031993               2062 bp    mRNA    linear   PRI 27-JUN-2005
DEFINITION  Homo sapiens corneodesmosin, mRNA (cDNA clone MGC:21340
            IMAGE:4747362), complete cds.
ACCESSION   BC031993
VERSION     BC031993.1  GI:21619291
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2062)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2062)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
```

```
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 28 Row: m Column: 21
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 26787969
            This clone grew slowly and was rescued by PCR.
FEATURES              Location/Qualifiers
     source           1. .2062
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:21340 IMAGE:4747362"
                      /tissue_type="Skin, squamous cell carcinoma"
                      /clone_lib="NCI_CGAP_Skn4"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6.1"
     gene             1. .2062
                      /gene="CDSN"
                      /note="synonyms: D6S586E, S, HTSS"
                      /db_xref="GeneID:1041"
                      /db_xref="MIM:602593"
     CDS              8. .1594
                      /gene="CDSN"
                      /codon_start=1
                      /product="corneodesmosin, precursor"
                      /protein_id="AAH31993.1"
                      /db_xref="GI:21619292"
                      /db_xref="GeneID:1041"
                      /db_xref="MIM:602593"
                      /translation="MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKD
                      PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                      AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSHSSSSSSF
                      QFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCS
                      SDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGG
                      LPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG
                      VPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSP
                      SSSRVPSSSSISGSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSK
                      SSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL
                      GPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match           97.7%;  Score 1553.2;  DB 8;  Length 2062;
  Best Local Similarity 99.0%;  Pred. No. 0;
  Matches 1574;  Conservative    0;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          8 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 67

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 127

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 187

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 247

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 307

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 367

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 427

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        428 TCGGGAAGCAGCGGCTCTCACTCGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 484

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 544

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 604

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 664

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        665 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 724

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 784

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        785 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 844

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        845 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 904

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        905 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 964

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        965 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1024

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1025 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1084

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1085 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1144

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1145 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1204

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||| ||||| |||||| ||||||||||||||||||||||||||||||||
Db       1205 AGTTCTAGCATTTCCGGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1264

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1265 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1324

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db       1325 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1384

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1385 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1444

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1445 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1504

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1505 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1564

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||
Db       1565 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1594


RESULT 13
BT019904
LOCUS       BT019904                1587 bp    mRNA    linear   SYN 28-OCT-2004
DEFINITION  Synthetic construct Homo sapiens corneodesmosin mRNA, partial cds.
ACCESSION   BT019904
VERSION     BT019904.1  GI:54696669
KEYWORDS    FLI_CDNA.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 1587)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Cloning of human full-length CDSs in BD Creator(TM) System Donor
            vector
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1587)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-OCT-2004) BD Biosciences Clontech, 1020 East Meadow
            circle, Palo Alto, California 94303, USA
COMMENT     This CDS clone is a part of a collection of human full length
            expression clones generated by BD Biosciences Clontech and the
            Harvard Institute of Proteomics. Each CDS has been cloned in two
            forms: with and without stop-codon (to allow fusion with C-terminal
            tag). The CDS has been  directionally cloned using BD In-Fusion(TM)
             cloning system between the SalI  and HindIII sites of the
            pDNR-DUAL  vector. Additional sequences in the clone:  'ACC' after
            SalI site  and before 'ATG' to provide Kozak consensus sequence;
            'GG' after  last codon and before HindIII site to maintain reading
            frame.  Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES             Location/Qualifiers
     source          1. .1587
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /clone="GH01206L1.0"
                     /clone_lib="BD Creator(TM) CDS Library derived from MGC
                     collection"
                     /lab_host="DH5alpha T1 resistant"
                     /note="Vector: pDNR-Dual"
     CDS             1. .>1587
                     /note="Mutations: 52:M->L;447:3DEL;597:OK"
                     /codon_start=1
                     /transl_table=11
                     /product="corneodesmosin"
                     /protein_id="AAV38707.1"
                     /db_xref="GI:54696670"
                     /translation="MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSHSSSSSSF
                     QFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCS
                     SDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGG
                     LPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG
                     VPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSP
                     SSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSK
                     SSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL
                     GPQLADPEVFLPQGELLDSPL"
ORIGIN

  Query Match             97.7%;  Score 1552.8;  DB 11;  Length 1587;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1573;  Conservative    0;  Mismatches   12;  Indels    3;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 477

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 537

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 597

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 657

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        658 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 717

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 777

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 837

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 897
```

```
Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        898 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 957

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1017

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1077

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1078 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1137

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1197

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||
Db       1198 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1257

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1258 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1317

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1318 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1377

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1437

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1497

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1557

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCAT 1588
              |||||||||||||||||||||||||||||
Db       1558 CCCCAAGGAGAGTTACTCGACAGTCCAT 1585
```

```
RESULT 14
AF224751
LOCUS       AF224751                1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.41 allele, partial
            cds.
ACCESSION   AF224751
VERSION     AF224751.1  GI:8699031
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.41"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.41"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.41"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78563.1"
                     /db_xref="GI:8699032"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             94.5%;  Score 1503.2;  DB 8;  Length 1764;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
               |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy         443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy         503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
               |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy         563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 642
```

```
Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1603 GTCCATAA 1610


RESULT 15
AF224753
LOCUS       AF224753               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.51 allele, partial
            cds.
ACCESSION   AF224753
VERSION     AF224753.1  GI:8699035
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.51"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.51"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.51"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78565.1"
                     /db_xref="GI:8699036"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQS
                     PCSPPGTGSPSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             94.5%;  Score 1503.2;  DB 8;  Length 1764;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342
```

```
Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
Db       343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
Db       403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
Db       463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
Db       523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
Db       583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
Db       643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
Db       703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
Db       763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy       803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
Db       823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
Db       883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
Db       943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
Db      1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
Db      1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
Db      1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
Db      1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1242

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
Db      1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
Db      1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
Db      1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
Db      1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
Db      1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
Db      1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy      1583 GTCCATAA 1590
Db      1603 GTCCATAA 1610
```

Search completed: March 22, 2006, 01:05:23
Job time : 8130 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2006, 22:50:05 ; Search time 6678 Seconds
(without alignments)
11139.782 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1583.6	99.6	1590	10	AY412935	AY412935 Homo sapi
	2	1123.4	70.7	1590	10	AY412936	AY412936 Pan trogl
	3	688.8	43.3	742	2	BG676174	BG676174 602622574
	4	641.4	40.3	1674	10	AY412937	AY412937 Mus muscu
	5	632.6	39.8	685	1	AU297843	AU297843 AU297843
	6	618	38.9	644	7	CV025058	CV025058 2636 Full
	7	615	38.7	673	1	AU134804	AU134804 AU134804
	8	592.2	37.2	648	5	BX505094	BX505094 DKFZp686G
	9	549.8	34.6	581	3	BP226265	BP226265 BP226265
	10	472.8	29.7	531	8	DN995832	DN995832 TC110851
	11	407.6	25.6	678	6	CF767753	CF767753 CES006300
	12	403.4	25.4	500	1	AU297870	AU297870 AU297870
	13	375	23.6	426	1	AU298085	AU298085 AU298085
	14	340.8	21.4	609	2	BB615758	BB615758 BB615758
	15	333.4	21.0	357	2	BF834417	BF834417 CM2-HT096
	16	326	20.5	371	2	BF835362	BF835362 CM2-HT096
	17	320.2	20.1	345	2	BF834425	BF834425 CM2-HT096
	18	319.8	20.1	354	2	BF834420	BF834420 CM2-HT096
c	19	284.4	17.9	541	2	BE140467	BE140467 IL0-HT000
c	20	269.2	16.9	710	10	AG319699	AG319699 Mus muscu
	21	264.2	16.6	545	2	BE030827	BE030827 129035 MA
	22	254.8	16.0	817	7	CK595258	CK595258 AGENCOURT
	23	244.2	15.4	617	7	CF906543	CF906543 A0448E01-
	24	233.2	14.7	605	1	AI746833	AI746833 ul08a01.y
c	25	231	14.5	1077	3	BM808976	BM808976 AGENCOURT
c	26	230	14.5	519	2	BE140469	BE140469 IL0-HT000
	27	226.4	14.2	591	2	BG384333	BG384333 303380 MA
	28	223.6	14.1	648	7	CF906181	CF906181 A0443H06-
	29	221	13.9	538	5	BQ560469	BQ560469 H4064E05-
c	30	220.8	13.9	696	10	AG319650	AG319650 Mus muscu
	31	220.6	13.9	567	1	AI151929	AI151929 ui47h05.y
	32	220.6	13.9	627	1	AI173863	AI173863 ue92e09.y
	33	220.2	13.8	510	6	CA568936	CA568936 K0435D05-
	34	220.2	13.8	517	6	CA562545	CA562545 K0302E03-
	35	220.2	13.8	523	6	CA563968	CA563968 K0320C05-
	36	219	13.8	591	1	AI317627	AI317627 uj21d06.y
	37	218.6	13.7	505	7	CF903218	CF903218 A0402E07-
	38	218.2	13.7	505	1	AI930552	AI930552 ul66b12.y
	39	216.8	13.6	541	6	CA566248	CA566248 K0349B04-
	40	216.6	13.6	482	7	CF906626	CF906626 A0449E07-
	41	216.6	13.6	496	6	CA562719	CA562719 K0304E10-
	42	216.6	13.6	497	6	CA562680	CA562680 K0304B03-
	43	214.4	13.5	662	7	CF909260	CF909260 A0534C04-
	44	214.4	13.5	672	7	CN662446	CN662446 A0734C04-
	45	213.8	13.4	467	5	BX922319	BX922319 BX922319

ALIGNMENTS

RESULT 1
AY412935
LOCUS AY412935 1590 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens CDSN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY412935
VERSION AY412935.1 GI:39768900
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1590)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1590)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .1590
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1. .>1590
/gene="CDSN"
/locus_tag="HCM4723"
ORIGIN

Query Match 99.6%; Score 1583.6; DB 10; Length 1590;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1586; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

```
RESULT 2
AY412936
LOCUS       AY412936               1590 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Pan troglodytes CDSN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412936
VERSION     AY412936.1  GI:39768901
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 1590)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 1590)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:9598"
     gene            <1..>1590
                     /gene="CDSN"
                     /locus_tag="HCM4723"
ORIGIN

  Query Match     .        70.7%;  Score 1123.4;  DB 10;  Length 1590;
  Best Local Similarity   71.1%;  Pred. No. 1.5e-265;
  Matches 1130;  Conservative    0;  Mismatches  460;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||| ||||||||||||||||||||||||||| |||||||||||||||||||||
Db          1 ATGGGCTCATCTCGGGCACCCTGGATGGGGCGTGTGGGCGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Db        121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Db        181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Db        241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Db        301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
                                                  ||||||||||||||||||
Db        361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              || |||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        421 TCAGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCNNNCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCNNCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              |||||||||||||||||||||||||||||     |||||||||||||||||||||||| |
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCNNNNNCCTGTGGTGGTGGTGGTGGACCAGNAT 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||
Db        781 GGTTCTGGTGCCCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Db        841 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
                              |||||||||||||||||||||||||||||||||||||||| |||
Db        901 NNNNNNNNNNNNNNNTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCNAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 3
BG676174
LOCUS       BG676174                742 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602622574F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4747362 5',
            mRNA sequence.
ACCESSION   BG676174
VERSION     BG676174.1  GI:13907570
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10597  row: d  column: 19
```

```
          High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .742
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4747362"
                     /tissue_type="squamous cell carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn4"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library  constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             43.3%;  Score 688.8;  DB 2;  Length 742;
  Best Local Similarity   98.4%;  Pred. No. 1.6e-158;
  Matches  728;  Conservative    0;  Mismatches    7;  Indels    5;  Gaps    3;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          8 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 67

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         68 CTGCTGGCTGGTCTCCTCCTGCCA-GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 126

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 186

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 246

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        247 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCA-GGT 305

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 365

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 425

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db        426 TCGGGAAGCAGCGGCTCTCACTCGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 482

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 542

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 602

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 662

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||| ||||||||||||||||||||||||||||||||| ||||||||||||||
Db        663 GACATCCCCGAATCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 722

Qy        721 CCCAGCTCCCACTCTGTGTC 740
              ||||||||||||||||||||
Db        723 CCCAGCTCCCACTCTGTGTC 742
```

```
RESULT 4
AY412937
LOCUS       AY412937               1674 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Mus musculus CDSN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412937
VERSION     AY412937.1  GI:39768902
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1674)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 1674)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .1674
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>1674
                     /gene="CDSN"
                     /locus_tag="HCM4723"
ORIGIN

  Query Match             40.3%;  Score 641.4;  DB 10;  Length 1674;
  Best Local Similarity   67.2%;  Pred. No. 9.8e-147;
  Matches 1141;  Conservative    0;  Mismatches  396;  Indels  162;  Gaps   10;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||| || || | ||||||| |||||||| |||||||| ||||||||| |||||||| ||
Db          1 ATGGGTTCCTCGCAGGCACCCCGGATGGGGAGTGTGGGAGGGCACGGGCTGATGGCATTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||| |||| |||||  | |||||||| | |||||||||||||||||| ||| |||| |||
Db         61 CTGATGGCCGGTCTTATTCTGCCAGGAATCTTGGCTAAGAGCATTGGGACCCTCTCGGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              |||||||||||||||||| | |||||||||| || ||||| || |||| ||| |||  |
Db        121 CCCTGTAAGGACCCCACGAGGATCACCTCCCCGAATGACCCTTGTCTCATTGGAAAGACT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              | ||||| | || |||| ||| |  |||| |||||||| |   |||||     |
Db        181 GGCTCCAACAGCATCAGCAGCCAAGGTGGATCCAGCAGCTTCAGCAGCCAAGGTGGATCC 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||  |                  | |||    | ||||||||||   |||||||
Db        241 AGCAGTTTCA-----------------GCAGCCATGGTGGATCCAGCAGCTCCCAGGGA 282

Qy        301 GGTTCTGCAGGATCTTT------TAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354
              ||||| ||||||| ||      ||| |||||||| ||||||||||||   |||||||
Db        283 AGTTCTTCAGGATCCTTGATATATAAACCAGGAACAGGGTATTCCCAGAGTAGCTACTCT 342

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
              |  || ||||||    | |  | |||   |  | ||  ||| |||   || |||||||||
Db        343 TATGGCTCTGGCGGCTCCCGGCCAGG---AGGCAGTGGCTCTCAGACAGGAAGCAGCAGC 399
```

```
Qy       415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGC--- 471
Db       400 TCTCAGACAGGAAGTAGTAGCTCTCAGTCAGGAAGCAGCGGCTCTCAGTCAGGAAGCAGT 459

Qy       472 ---------------AGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAAT 516
Db       460 GGCTCTCAGTCAGGAAGCAGCGGCTCCCAGTCAGGAAGCAGCGGCTCCCAGTCAGGAAGC 519

Qy       517 GGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAA------TACTAAACCCTTCC 570
Db       520 AGCGGCTCCCAGTCAGGAAGCAGCGGCTCTCAGTCAGGAAGATGGGTAAGCAGCAGTTCT 579

Qy       571 CAGCCTGGACAAAGCTCTTCCTCTTCCCAAA---------------------------- 601
Db       580 CAATGGGTAAGCAGCAGTTCTCAGTCCGGAAGTTCAGGAAGCAGTCGCGACAGACCCGGA 639

Qy       602 ---------------------------------------------CCTCTGGGGTATCC 615
Db       640 AGTGGCTCTGCTCTACCAACCGGTGACAAGACATCAGGCATGTCTCAGTCTGGAGGCTCG 699

Qy       616 AGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCT 675
Db       700 TCCACTTCCCAGAGCAGCAGCTCCAACCTGCGGCCCTGTAGCTCCAACGTGCCCGACTCT 759

Qy       676 CCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCT 735
Db       760 CCCTGCAGCGGGGGGCCTGTCATCACACACTCGGGGCCCTACATCTCTGGCACCCACACC 819

Qy       736 GTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCT 795
Db       820 GTATCCGGTGGCCAGAGACCTGTAGTGGTGGTAGTGGAGCAACATGGCTCTGGTGGCCCT 879

Qy       796 GGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCA 855
Db       880 GG---GTTTCAAGGCATGCCCTGTAGCAATGGAGGCCCCGCAGGTAAGCCCTGCCCTCCC 936

Qy       856 ATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGT 915
Db       937 ATCACCTCTGTCCAGAAACCCTATGGCGGCTATGAGGTGGTGGGTGGCTCTGCCAACAGT 996

Qy       916 TATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACC 975
Db       997 TATCTGGTCCCAGGCATGACCTACAGTGGGGGTAAAATCTACCCCGTGGGCTACTTCACC 1056

Qy       976 AAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATC 1035
Db      1057 AAGGACAACCCTATCAGGGGCTCGCCAGGGGCCCCCTCCTTTGCAGCTGGGCCCCCAGTC 1116

Qy      1036 TCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC- 1094
Db      1117 TCTGAGGGCAAGTACTTCTCTAGCAACCCCATCATCCCCAGCCGTGGCTCTTCCAGTTCC 1176

Qy      1095 ---------------GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db      1177 AGTGGCTACCCAGTGGGCGTTGTGTTCCAGCCCGTGGGCTCTGGCGGGGTTCAGCCCTGT 1236

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db      1237 GGCACTGGCTCTGTCAGCTCTAAGGGCCCTTGCTC---------------GGGTACGAGA 1281

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTG---CTTCC 1257
Db      1282 ATTCAGATCACCTCCAGCAGCTCCAGTACCTCCTACCACCCCTGCAGCGGTGGTTCTTCC 1341

Qy      1258 CAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCG 1317
Db      1342 CAGGGGCCTTGCTCCTCACCAGGCACCGGCTCCATCAGTGGGGGAAGCTCCTCCCTCTCC 1401

Qy      1318 AGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377
Db      1402 TCTGGCAAAATCGTCCTTCAGCCCTGCGGCAGCAAATCTACCTCTTCGGGTTACCCCTGC 1461

Qy      1378 ATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCC 1437
Db      1462 CTTTCTGTTCCCCTCCTCCCCATTGAATGGTGGTCTGAATGGCTCTCCTCAGCCTGTCCCC 1521

Qy      1438 TCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATC 1497
Db      1522 TCCGTGGGTGTCAAGCTCTGTGGCCTCAACAGCCCTGGAAGAGTGCCCTGCCGTTCCATC 1581

Qy      1498 CGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTC 1557
Db      1582 CGCAACATTCTGACCCAAGTGAAGCCTTTGGGGCCCCAGTTAATGGATCCCAAAGTTTCT 1641

Qy      1558 CTACCCCAAGGAGAGTTAC 1576
Db      1642 CTGCCACAGGGAGAGCCAC 1660
```

RESULT 5
AU297843
LOCUS AU297843 685 bp mRNA linear EST 08-MAY-2003
DEFINITION AU297843 male adult skin, full-length enriched chimpanzee cDNA library Pan troglodytes verus cDNA clone PstA0073 5' similar to human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION AU297843
VERSION AU297843.1 GI:29532142
KEYWORDS EST.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 685)
AUTHORS Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N., Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
TITLE Analysis of 5'-end sequences of chimpanzee cDNAs
JOURNAL Genome Res. 13 (5), 1022-1026 (2003)
PUBMED 12727913
COMMENT Contact: Momoki Hirai
Department of Integrated Biosciences
The University of Tokyo, Graduate School of Frontier Sciences
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
Tel: 81-4-7136-3688
Fax: 81-4-7136-3687
Email: mhirai@k.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .685
/organism="Pan troglodytes verus"
/mol_type="mRNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone="PstA0073"
/sex="male"
/tissue_type="skin"
/dev_stage="adult"
/clone_lib="male adult skin, full-length enriched chimpanzee cDNA library"
ORIGIN

Query Match 39.8%; Score 632.6; DB 1; Length 685;
Best Local Similarity 98.5%; Pred. No. 1.1e-144;
Matches 649; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

```
Qy         1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db        28 ATGGGCTCATCTCGGGCACCCTGGATGGGGCGTGTGGGCGGGCACGGGATGATGGCACTG 87

Qy        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db        88 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 147

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db       148 CCCTGTAAGGACCCCATGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 207
```

```
Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 267

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        268 GCCAGAAGCTCTGGTGGTGGCTCCAGCGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 327

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 387

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 447

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TCAGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 507

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 567

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||| |||||||||
Db        568 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTT-CTCTTCCCAG 626

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTC 659
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTC 685
```

RESULT 6
CV025058

```
LOCUS       CV025058                644 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  2636 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC031993, mRNA sequence.
ACCESSION   CV025058
VERSION     CV025058.1  GI:51482951
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGGCTCGTCTCGGGCA
            BACKWARD: TATGGACTGTCGAGTAACTCTCCTT
            Insert Length: 644    Std Error: 36.00
            Plate: 11068  row: 04  column: B
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 98
            High quality sequence stop: 643
            POLYA=No.
FEATURES            Location/Qualifiers
     source         1. .644
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /tissue_type="mixed"
                    /clone_lib="Full Length cDNA from the Mammalian Gene
                    Collection"
                    /note="Vector: mixed; The ORFs were PCR amplified from the
                    MGC (Mammalian Gene Collection) as of April 2004 and
                    cloned by recombinational Gateway cloning into pDONR223
                    Donor vector. Reference : MGC (Mammalian Gene Collection)
                    Program Team, Generation and Initial Analysis of more than
                    15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                    2002, 99(26), 16899-16903"
ORIGIN
```

```
  Query Match             38.9%;  Score 618;  DB 7;  Length 644;
  Best Local Similarity   98.8%;  Pred. No. 4.5e-141;
  Matches  634;  Conservative    0;  Mismatches    5;  Indels    3;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 477

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 537

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||| |||||||||
Db        538 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 597

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAAC 642
              |||| ||||||||||||||||||||||||||||||||||||
Db        598 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAAC 639
```

RESULT 7
AU134804

```
LOCUS       AU134804                673 bp    mRNA    linear   EST 02-AUG-2002
DEFINITION  AU134804 PLACE1 Homo sapiens cDNA clone PLACE1000541 5', mRNA
            sequence.
ACCESSION   AU134804
```

```
VERSION     AU134804.1  GI:10995343
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .673
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1000541"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             38.7%;  Score 615;  DB 1;  Length 673;
  Best Local Similarity   96.9%;  Pred. No. 2.5e-140;
  Matches  624;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 88

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 148

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 208

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 268

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 328

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 388

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 448

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 508

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 568

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||| ||||||||||||||||| ||||||||||||||||||||||||||
Db        569 AACTCTTACCGCGGGATACTAAACCCTTCCCAACCTGGACAAAGCTCTTCCTCTTCCCAA 628

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCA 644
              | || |||||||||| |      ||| |||| || |    ||| ||
Db        629 ANCTTTGGGGTATCAACANTNGGCAAAAGNGTTANGCTCAAACA 672


RESULT 8
BX505094
LOCUS       BX505094                 648 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686G12192_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686G12192 5', mRNA sequence.
ACCESSION   BX505094
VERSION     BX505094.1  GI:32032777
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No sl sequence available.
            This clone (DKFZp686G12192) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686G12192"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             37.2%;  Score 592.2;  DB 5;  Length 648;
  Best Local Similarity   99.2%;  Pred. No. 1e-134;
  Matches  616;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 88

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 148

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 208
```

```
Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 268

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 328

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 388

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 448

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 508

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 568

Qy        541 AACTCTTAGCGCGG-AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 599
              |||||||| |||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        569 AACTCTTACCGCGGNAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTT-CCA 627

Qy        600 AACCTCTGGGGTATCCAGCAG 620
              ||||| ||||||||||||||||
Db        628 AACCTTTGGGGTATCCAGCAG 648
```

```
RESULT 9
BP226265
LOCUS       BP226265                581 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP226265 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
            clone DMC02127, mRNA sequence.
ACCESSION   BP226265
VERSION     BP226265.1  GI:52099170
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DMC02127"
                     /clone_lib="Sugano cDNA library, dermoid cancer"
                     /note="dermoid cancer"
ORIGIN

  Query Match             34.6%;  Score 549.8;  DB 3;  Length 581;
  Best Local Similarity   99.6%;  Pred. No. 2.8e-124;
  Matches  551;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCT 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 GGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCT 88

Qy         63 GCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 148

Qy        123 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGA 208

Qy        183 CTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGC 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGC 268

Qy        243 CAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGG 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGG 328

Qy        303 TTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATC 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATC 388

Qy        363 TGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTC 422
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTC 448

Qy        423 GGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTT 508

Qy        483 TCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAA 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAA 568

Qy        543 CTCTTAGCGCGGA 555
              |||||| ||||||
Db        569 CTCTTACCGCGGA 581
```

```
RESULT 10
DN995832
LOCUS       DN995832                531 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC110851 Human breast cancer tissue, large insert, pCMV expression
            library Homo sapiens cDNA clone TC110851 5' similar to Homo sapiens
            corneodesmosin (CDSN), mRNA sequence.
ACCESSION   DN995832
VERSION     DN995832.1  GI:66255659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
            6 Taft Ct. Suite 100
```

```
            Rockville, MD 20850
            Tel: (301) 340-3188
            http://www.origene.com
            Seq primer: pCMV6 5prime forward vector primer, OriGene
            Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC110851"
                     /tissue_type="Breast cancer"
                     /clone_lib="Human breast cancer tissue, large insert, pCMV
                     expression library"
                     /note="Organ: Mammary gland (cancer tissue); Vector:
                     pCMV6-XL5; Site_1: EcoR1; Site_2: Xho1/Sal1 compatible end
                     ligatio; Oligo-dT primed reverse transcription optimized
                     for large and GC rich mRNA transcripts, cDNA size
                     selection, optimized ligation for large inserts into
                     mammalian expression vector, random clones selected for
                     end sequence verification of full-length genes"
ORIGIN

  Query Match             29.7%;  Score 472.8;  DB 8;  Length 531;
  Best Local Similarity   99.4%;  Pred. No. 2.6e-105;
  Matches  474;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         55 ATGGGCTCGTCTCGGGCACCCTGAATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 114

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 174

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 234

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 294

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 354

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        355 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGNGTATTCCCAGGTCAGCTACTCCTCCGGA 414

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 474

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGC 477
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGC 531


RESULT 11
CF767753
LOCUS       CF767753                678 bp    mRNA    linear   EST 17-OCT-2003
DEFINITION  CES006300 Bos taurus skin cDNA library Bos taurus cDNA clone
            CCL006300 5', mRNA sequence.
ACCESSION   CF767753
VERSION     CF767753.1  GI:37716972
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Wang,Y.H., McWilliam,S. and Lehnert,S.
  TITLE     Transcription profiling of cattle skin
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dr Yonghong Wang
            Functional Genomics Lab
            CSIRO Livestock Industries
            Level 5, Queensland Biosciences Precinct, University of Queensland,
            306 Carmody Road St.Lucia QLD Australia
            Tel: 07 3214 2445
            Fax: 07 3214 2685
            Email: Yonghong.Wang@csiro.au
            Plate: 72  row: F  column: 03.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Hereford Shorthorn"
                     /db_xref="taxon:9913"
                     /clone="CCL006300"
                     /sex="female"
                     /tissue_type="pooled"
                     /dev_stage="Adult"
                     /lab_host="XL1-BlueMRF'strain"
                     /clone_lib="Bos taurus skin cDNA library"
                     /note="Organ: skin; Vector: Uni-ZAPXR; Site_1: EcoRI;
                     Site_2: Xho I; Library made from pooled skin of adult
                     female Hereford-Shorthorn."
ORIGIN

  Query Match             25.6%;  Score 407.6;  DB 6;  Length 678;
  Best Local Similarity   81.7%;  Pred. No. 3.3e-89;
  Matches  495;  Conservative    0;  Mismatches   84;  Indels   27;  Gaps    1;

Qy       1012 TCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATC 1071
              |||||||| ||||||||||||||||||||||||| ||||||||  ||||||||||||||
Db          2 TCCTTTGCTGCTGGGCCCCCCATCTCTGAGGGCAAGTACTTCTCTGGCAACCCCATCATC 61

Qy       1072 CCCAGCCAG--------------------------TCGGCAGCTTCCTCGGCCATTGCG 1104
              ||||||||                          || | |||||||||| | | | |
Db         62 CCCAGCCACAGCTCTTCTAGCTCCAATATCTACCAATCCGGAGCTTCCTCGACTGTGGTG 121

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
              |||||||||||||  |||||||||||||| |||||  | |||||| ||||||| |||
Db        122 TTCCAGCCAGTGGGCTCTGGTGGGGTCCAGCCCTGTGTGGTCGGCTCCTCGGGCTCTAAG 181

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
              || |||||||| | |||||| ||| ||||| ||||||| || ||||||||||||  |
Db        182 GGGCCCTGCTCCCTCTCCAGCTCTGGAGTCTCCAGCAGGTCCAGCATTTCCAGCAGTTCT 241

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
              ||||| | | ||||||||| ||  ||| ||| ||| | ||||||||||||||||||||
Db        242 AGTTCATCTTTCCATCCCTGTGGTGGTGTTTCTCAGGGGCCCTGCTCCCCACCAGGCACT 301

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
              |||||| ||||| ||||||| || |||| |||  |||||||||||||||||||| |||
Db        302 GGCTCCCTCAGCGGCAGCTCTAGCTCCCTATCTGGTGGCAAAATCATCCTTCAGCCCTGT 361

Qy       1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
              |||||||||||||| || |||||||||||| || ||||| ||||||| |||||||| |
Db        362 GGCAGCAAGTCCAGCGCTCCTGGTCACCCTTGTATTTCTGTTTCCTCCTCGACACTGAGT 421

Qy       1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
              ||||| ||  |||||||||||||| ||||| ||||| |||||||||||||||||||  ||
Db        422 GGGGGTCCAAATGGCTCTCCCCAACCTGACCCCTCAGCTGGTGCCAAGCCCTGTGGTGCC 481

Qy       1465 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1524
              |||     |||||||| ||| || |||||||||||||| |||||||||| |||||||||
Db        482 GGCAACTATGGAAAGACCCCTTGTCGCTCCATCCGGGACATCCTAGCCCAGGTGAAGCCT 541

Qy       1525 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1584
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||  |||  |||||
Db        542 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAGGGAGATCTACCTAACAGT 601
```

```
Qy       1585 CCATAA 1590
              | |||
Db        602 TCTTAA 607


RESULT 12
AU297870
LOCUS       AU297870                500 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  AU297870 male adult skin, full-length enriched chimpanzee cDNA
            library Pan troglodytes verus cDNA clone PstA0195 5' similar to
            human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION   AU297870
VERSION     AU297870.1  GI:29532169
KEYWORDS    EST.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N.,
            Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
  TITLE     Analysis of 5'-end sequences of chimpanzee cDNAs
  JOURNAL   Genome Res. 13 (5), 1022-1026 (2003)
   PUBMED   12727913
COMMENT     Contact: Momoki Hirai
            Department of Integrated Biosciences
            The University of Tokyo, Graduate School of Frontier Sciences
            5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
            Tel: 81-4-7136-3688
            Fax: 81-4-7136-3687
            Email: mhirai@k.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .500
                     /organism="Pan troglodytes verus"
                     /mol_type="mRNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone="PstA0195"
                     /sex="male"
                     /tissue_type="skin"
                     /dev_stage="adult"
                     /clone_lib="male adult skin, full-length enriched
                     chimpanzee cDNA library"
ORIGIN

  Query Match             25.4%;  Score 403.4;  DB 1;  Length 500;
  Best Local Similarity   98.1%;  Pred. No. 3.3e-88;
  Matches  407;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db         28 ATGGGCTCATCTCGGGCACCCTGGATGGGGCGTGTGGGCGGGCACGGGATGATGGCACTG 87

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         88 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTAGCACCTTCTCAGAC 147

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        148 CCCTGTAAGGACCCCATGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 207

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        208 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCANTTCTGGCAGCTCCATTTCCAGT 267

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        268 GCCAGAAGCTCTGGTGGTGGCTCCAGCGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 327

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GGTTCTGNAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 387

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCT 415
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCT 442


RESULT 13
AU298085
LOCUS       AU298085                426 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  AU298085 male adult skin, full-length enriched chimpanzee cDNA
            library Pan troglodytes verus cDNA clone PstA2015 5' similar to
            human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION   AU298085
VERSION     AU298085.1  GI:29532384
KEYWORDS    EST.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N.,
            Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
  TITLE     Analysis of 5'-end sequences of chimpanzee cDNAs
  JOURNAL   Genome Res. 13 (5), 1022-1026 (2003)
   PUBMED   12727913
COMMENT     Contact: Momoki Hirai
            Department of Integrated Biosciences
            The University of Tokyo, Graduate School of Frontier Sciences
            5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
            Tel: 81-4-7136-3688
            Fax: 81-4-7136-3687
            Email: mhirai@k.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Pan troglodytes verus"
                     /mol_type="mRNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone="PstA2015"
                     /sex="male"
                     /tissue_type="skin"
                     /dev_stage="adult"
                     /clone_lib="male adult skin, full-length enriched
                     chimpanzee cDNA library"
ORIGIN

  Query Match             23.6%;  Score 375;  DB 1;  Length 426;
  Best Local Similarity   97.5%;  Pred. No. 3.1e-81;
  Matches  389;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db         28 ATGGGCTCATCTCGGGCACCCTGGATGGGGCGTGTGGGCGGGCACGGGATGATGGCACTG 87

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         88 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 147

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        148 CCCTGTAAGGACCCCATGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 207

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        208 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 267

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        268 GCCAGAAGCTCTGGTGGTGGCTCCAGCGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 327

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGT-CAGCTACTCCTCCGG 359
```

```
                ||||||||||||||||||||||||||||||||||||||||||| || |  ||||| ||||||
Db        328 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCANGTCCNNCTACTNCTCCGG 387

Qy        360 ATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCA 398
                ||||||||||| ||||||||||||||||||||||||||
Db        388 ATCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCA 426
```

RESULT 14
BB615758
LOCUS BB615758 609 bp mRNA linear EST 26-OCT-2001
DEFINITION BB615758 RIKEN full-length enriched, adult male testis Mus musculus cDNA clone 4930556G08 5', mRNA sequence.
ACCESSION BB615758
VERSION BB615758.1 GI:16455990
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 609)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .609
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4930556G08"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male testis"
/note="Site_1: XhoI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "
ORIGIN

Query Match 21.4%; Score 340.8; DB 2; Length 609;
Best Local Similarity 76.9%; Pred. No. 9.2e-73;
Matches 449; Conservative 0; Mismatches 117; Indels 18; Gaps 2;

```
Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
              |||||  |  ||    || |  ||| |||  ||||||||||| ||| |||||||| ||  ||
Db         29 TCTGGAGGCTCGTCCACTTCCCAGAGCAGCAGCTCCAACCTGCGGCCCTGTAGCTCCAAC 88

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
              | |||||||||||||||||| || |||||  || || | ||||| || |||||||||| |
Db         89 GTGCCCGACTCTCCCTGCAGCGGGGGGCCTGTCATCACACACTCGGGGCCCTACATCTCT 148

Qy        724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
              || ||||| |||| || || || ||||| ||||| |||||||||||||| || || ||
Db        149 GGCACCCACACTGTATCCGGTGGCCAGAGACCTGTAGTGGTGGTGGTGGAGCAACATGGC 208

Qy        784 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
              ||||||| ||||||   | ||||||||    |||||||||||||| ||||   |||| |||
Db        209 TCTGGTGGCCCTGG---GTTTCAAGGCATGCCCTGTAGCAATGGAGGCCCCGCAGGTAAG 265

Qy        844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
              ||||| || || |||||||||||  | ||| |||||||  ||||| ||||||||||||||
Db        266 CCCTGCCCTCCCATCACCTCTGTCCAGAAACCCTATGGCGGCTATGAGGTGGTGGGTGGC 325

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
              ||  |  |||||||||||||| ||||||||||||||||||  |||||||||||| || |||
Db        326 TCTGCCAACAGTTATCTGGTCCCAGGCATGACCTACAGTGGGGGTAAAATCTACCCCGTG 385

Qy        964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
              |||||||||||||| || |||||| | |  ||||| |||||||  ||| ||||||||||||
Db        386 GGCTACTTCACCAAGGACAACCCTATCAGGGGCTCGCCAGGGGCCCCCTCCTTTGCAGCT 445

Qy       1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
              ||||||||  |||||||||||||| |||||||| ||||||||||||||||||||||||
Db        446 GGGCCCCCAGTCTCTGAGGGCAAGTACTTCTCTAGCAACCCCATCATCCCCAGCCGTGGC 505

Qy       1084 GCAGCTTCCTC----------------GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
                |  |    ||                || | ||||||||||||| |||||  |||| |||
Db        506 TCTTCCAGTTCCAGTGGCTACCCAGTGGGCGTTGCGTTCCAGCCCGTGGGCTCTGGCGGG 565

Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTG 1172
              || |||| ||||||    |||||     |||| ||||| || ||
Db        566 GTTCAGCCCTGTGGCACTGGCTCTGTCAGCTCTAAGGGCCCTTG 609
```

RESULT 15
BF834417
LOCUS BF834417 357 bp mRNA linear EST 13-JAN-2001
DEFINITION CM2-HT0968-111100-500-b11 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF834417
VERSION BF834417.1 GI:12184792
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 357)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-111100-500-b11&t3=2000-11-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 357.

FEATURES Location/Qualifiers
source 1. .357
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0968"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 21.0%; Score 333.4; DB 2; Length 357;
Best Local Similarity 99.7%; Pred. No. 5.3e-71;
Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        872 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 82

Qy        932 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 142

Qy        992 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 1051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 202

Qy       1052 TCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGC 1111
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        203 TCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCGGCTTCCTCGGCCATTGCGTTCCAGC 262

Qy       1112 CAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCT 1171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCT 322

Qy       1172 GCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 1206
              |||||||||||||||||||||||||||||||||||
Db        323 GCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 357
```

Search completed: March 22, 2006, 02:56:46
Job time : 6683 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2006, 22:33:00 ; Search time 1001 Seconds
(without alignments)
10586.275 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1586.8	99.8	1590	2	AAV63614	Aav63614 DNA seque
2	1586.8	99.8	1590	5	AAS43749	Aas43749 DNA encod
3	1586.8	99.8	1725	9	ABT43850	Abt43850 DNA of th
4	1586.8	99.8	1725	12	ADN04225	Adn04225 Antipsori
5	1585.2	99.7	2057	12	ADP10483	Adp10483 Reference
6	1582	99.5	1590	2	AAV63541	Aav63541 DNA seque
7	1580	99.4	1725	4	AAH73624	Aah73624 DNA seque
8	1570.8	98.8	2284	3	AAZ50582	Aaz50582 Human epi
9	1570	98.7	2530	6	ABL68811	Abl68811 Kidney ca
10	1570	98.7	2530	9	ABT43848	Abt43848 DNA of th
11	1570	98.7	2530	12	ADQ21503	Adq21503 Human sof
12	1503.2	94.5	4740	9	ABT43849	Abt43849 Partial D
13	1503.2	94.5	5337	4	AAH45309	Aah45309 Human MHC
14	1503.2	94.5	25235	4	AAH45310	Aah45310 Human SEE
15	1503.2	94.5	55050	6	ABQ75680	Abq75680 Human SEE
16	1501.2	94.4	1506	9	ABT43852	Abt43852 Exon 2 of
17	1411.4	88.8	2551	12	ADQ25340	Adq25340 Human sof
18	655.8	41.2	659	12	ACH92172	Ach92172 Human gen
19	505	31.8	505	12	ACH78468	Ach78468 Human gen
20	85	5.3	85	9	ABT43851	Abt43851 Exon 1 of
21	84.6	5.3	85	5	AAS43706	Aas43706 Corneodes
22	84.2	5.3	85	5	AAS43705	Aas43705 Corneodes
23	80.6	5.1	81	5	AAS43709	Aas43709 Corneodes
24	80.6	5.1	81	5	AAS43700	Aas43700 Corneodes
25	80.6	5.1	81	6	ABQ75645	Abq75645 Human SEE
26	79.6	5.0	80	5	AAS43703	Aas43703 Corneodes
27	79	5.0	79	5	AAS43704	Aas43704 Corneodes
28	78.6	4.9	79	5	AAS43708	Aas43708 Corneodes
29	78.6	4.9	79	5	AAS43702	Aas43702 Corneodes
30	74.8	4.7	76	6	ABQ75642	Abq75642 Human SEE
31	74.6	4.7	75	5	AAS43710	Aas43710 Corneodes
32	74.6	4.7	75	6	ABQ75643	Abq75643 Human SEE
33	74	4.7	76	5	AAS43711	Aas43711 Corneodes
34	73.6	4.6	74	5	AAS43715	Aas43715 Corneodes
35	73.2	4.6	74	6	ABQ75639	Abq75639 Human SEE
36	71.6	4.5	72	5	AAS43716	Aas43716 Corneodes
37	71.6	4.5	72	5	AAS43707	Aas43707 Corneodes
38	71.2	4.5	72	6	ABQ75638	Abq75638 Human SEE
39	70.6	4.4	71	5	AAS43686	Aas43686 Corneodes
40	70.6	4.4	71	6	ABQ75644	Abq75644 Human SEE
41	68.6	4.3	69	5	AAS43712	Aas43712 Corneodes
42	67.8	4.3	69	6	ABQ75641	Abq75641 Human SEE
43	65.8	4.1	67	6	ABQ75640	Abq75640 Human SEE
44	65	4.1	67	5	AAS43714	Aas43714 Corneodes
45	64.6	4.1	65	5	AAS43701	Aas43701 Corneodes

ALIGNMENTS

RESULT 1
AAV63614
ID AAV63614 standard; DNA; 1590 BP.
XX
AC AAV63614;
XX
DT 10-FEB-1999 (first entry)
XX
DE DNA sequence encoding a protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1590
FT /*tag= a
XX
PN FR2761362-A1.
XX
PD 02-OCT-1998.
XX
PF 28-MAR-1997; 97FR-00003899.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586093/50.
DR P-PSDB; AAW80498.
XX
PT New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT or therapeutic treatment of trophic disorders or thinning of the
PT epidermis.
XX
PS Claim 27; Page 32; 35pp; French.
XX

CC The present sequence encodes a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products
XX
SQ Sequence 1590 BP; 291 A; 534 C; 419 G; 346 T; 0 U; 0 Other;

Query Match 99.8%; Score 1586.8; DB 2; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
AAS43749
ID AAS43749 standard; DNA; 1590 BP.
XX
AC AAS43749;
XX
DT 18-DEC-2001 (first entry)
XX
DE DNA encoding corneodesmosin.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.

XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
DR P-PSDB; AAU27550.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Claim 1; Fig 1; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 1590 BP; 291 A; 534 C; 419 G; 346 T; 0 U; 0 Other;

Query Match 99.8%; Score 1586.8; DB 5; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
```

```
Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 3
ABT43850
ID   ABT43850 standard; DNA; 1725 BP.
XX
AC   ABT43850;
XX
DT   06-NOV-2003  (first entry)
XX
DE   DNA of the human corneodesmosin (CDSN) mRNA.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW   gene.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             15. .1604
FT                   /*tag=  a
FT                   /product= "Human CDSN protein (corneodesmosin)"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   P-PSDB; ABJ39148.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Example 1; Fig 2d; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polynucleotide sequence is the DNA copy of the mRNA that encodes the
CC   human corneodesmosin (CDSN) protein of the invention
XX
SQ   Sequence 1725 BP; 324 A; 570 C; 454 G; 377 T; 0 U; 0 Other;

  Query Match             99.8%;  Score 1586.8;  DB 9;  Length 1725;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1588;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
               |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 4
ADN04225
ID ADN04225 standard; cDNA; 1725 BP.
XX
AC ADN04225;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #312.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR P-PSDB; ADN04226.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 619; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 1725 BP; 324 A; 570 C; 454 G; 377 T; 0 U; 0 Other;

  Query Match   99.8%; Score 1586.8; DB 12; Length 1725;
  Best Local Similarity   99.9%; Pred. No. 0;
  Matches 1588; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974
```

```
Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 5
ADP10483
ID ADP10483 standard; DNA; 2057 BP.
XX
AC ADP10483;
XX
DT 12-AUG-2004 (first entry)
XX
DE Reference mRNA sequences for marker probe #160.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 80; SEQ ID NO 492; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.
XX
SQ Sequence 2057 BP; 416 A; 681 C; 490 G; 470 T; 0 U; 0 Other;

```
  Query Match            99.7%;  Score 1585.2;  DB 12;  Length 2057;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1587;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
```

```
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604


RESULT 6
AAV63541
ID   AAV63541 standard; DNA; 1590 BP.
XX
AC   AAV63541;
XX
DT   05-FEB-1999  (first entry)
XX
DE   DNA sequence encoding a protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1590
FT                   /*tag=  a
XX
PN   FR2761363-A1.
XX
PD   02-OCT-1998.
XX
PF   11-SEP-1997;   97FR-00011317.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586094/50.
DR   P-PSDB; AAW85055.
XX
PT   New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT   cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT   epidermis etc., also compositions containing protease that degrades
PT   corneodesmosin.
XX
PS   Claim 27; Page 33; 35pp; French.
XX
CC   The present sequence encodes a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 1590 BP; 291 A; 534 C; 418 G; 347 T; 0 U; 0 Other;

  Query Match             99.5%;  Score 1582;  DB 2;  Length 1590;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1585;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTATGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 7
AAH73624
ID AAH73624 standard; DNA; 1725 BP.
XX
AC AAH73624;
XX
DT 01-OCT-2001 (first entry)
XX
DE DNA sequence #1 for vulgar psoriasis diagnosis.
XX
KW MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism; ds.
XX
OS Unidentified.
XX
PN WO200142458-A1.
XX
PD 14-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-JP008624.
XX
PR 06-DEC-1999; 99JP-00346867.
XX
PA (INOK/) INOKO H.
XX
PI Inoko H, Tamiya G;
XX
DR WPI; 2001-381680/40.
DR P-PSDB; AAG64152.
XX
PT New primer DNA, useful for detecting vulgar psoriasis.
XX
PS Disclosure; Fig 2; 106pp; Japanese.
XX
CC The invention relates to a method of diagnosing vulgar psoriasis using
CC primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC By analysing the sequences of these genes in Japanese patients with
CC psoriasis and in normal subjects, it has been found that some of the
CC examined polymorphisms correlate significantly to the group of patients
CC with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC these gene polymorphisms. The present sequence is provided in the
CC specification
XX
SQ Sequence 1725 BP; 322 A; 566 C; 449 G; 377 T; 0 U; 11 Other;

  Query Match   99.4%; Score 1580; DB 4; Length 1725;
  Best Local Similarity 99.2%; Pred. No. 0;
  Matches 1577; Conservative   9; Mismatches   4; Indels   0; Gaps   0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74
Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
```

```
Db        75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            |||||||:|||||||||||||||||||||| ||||:|||||||||||||||||||||||||
Db       435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            :|||||||||||||||||||:|||:||:|||||||||||||||||||||||||||||||||
Db      1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            ||||||||||||||||||:||||||||||
Db      1575 CCCCAAGGAGAGTTACTCRACAGTCCATAA 1604
```

```
RESULT 8
AAZ50582
ID   AAZ50582 standard; cDNA; 2284 BP.
XX
AC   AAZ50582;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human epidermal protein-6 cDNA.
XX
KW   Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW   dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW   arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1. .142
FT                   /*tag=  b
FT   CDS             47. .1636
FT                   /*tag=  a
FT                   /product= "Human epidermal protein-6"
FT   mat_peptide     148. .1636
FT                   /*tag=  c
FT                   /product= "Mature human epidermal protein-6"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
```

```
DR   WPI; 2000-195295/17.
DR   P-PSDB; AAY44989.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 7; Page 78-79; 82pp; English.
XX
CC   The present cDNA sequence encodes human epidermal protein-6 (HEPI). The
CC   cDNA clone is derived from PTHYNOT03 library which was constructed using
CC   RNA isolated from left parathyroid tissue of a 69-year-old caucasian
CC   female during a partial parathyroidectomy. Recombinant vectors comprising
CC   HEPI cDNA are introduced into host cells for protein expression. The HEPI
CC   proteins are useful for the treatment of epithelial disorders, including
CC   dyshidrotic eczema and scabies, cell proliferative disorders including
CC   actinic keratosis and arteriosclerosis, and autoimmune/inflammatory
CC   disorders like acquired immune deficiency syndrome (AIDS) and Addison's
CC   disease. Pharmaceutical compositions comprising HEPI proteins are useful
CC   for treating disorders associated with altered HEPI expression
XX
SQ   Sequence 2284 BP; 452 A; 775 C; 534 G; 523 T; 0 U; 0 Other;

  Query Match             98.8%;  Score 1570.8;  DB 3;  Length 2284;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1578;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||
Db       1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636


RESULT 9
ABL68811
ID   ABL68811 standard; DNA; 2530 BP.
XX
AC   ABL68811;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Kidney cancer related gene sequence SEQ ID NO:7148.
XX
```

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 7148; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2530 BP; 496 A; 840 C; 593 G; 601 T; 0 U; 0 Other;

Query Match 98.7%; Score 1570; DB 6; Length 2530;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
```

```
                |||||||
```

```
Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            |||||||||||||||||||||||||||||
Db      1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

RESULT 11
ADQ21503
ID ADQ21503 standard; DNA; 2530 BP.
XX
AC ADQ21503;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 4323.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4323; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 2530 BP; 496 A; 840 C; 593 G; 601 T; 0 U; 0 Other;

  Query Match   98.7%; Score 1570; DB 12; Length 2530;
  Best Local Similarity   99.6%; Pred. No. 0;
  Matches 1584; Conservative   0; Mismatches   5; Indels   1; Gaps   1;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74
```

```
Qy        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             |||||||||||||||||||||||||||||
Db      1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

RESULT 12
ABT43849
ID ABT43849 standard; DNA; 4740 BP.
XX
AC ABT43849;
XX
DT 06-NOV-2003 (first entry)
XX
DE Partial DNA of the MHC region containing the human corneodesmosin gene.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW gene; MHC.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	exon	139. .237
FT		/*tag= b
FT		/number= 1
FT		/note= "This sequence is also found in ABT43851"
FT	CDS	153. .4546
FT		/*tag= a
FT		/product= "Human CDSN protein (corneodesmosin)"
FT		/transl_except= (pos: 234. .237; aa: Pro)
FT		/note= "This codon has an apparent one nucleotide
FT		insertion that alters the reading frame"
FT		/transl_except= (pos: 4178. .4180; aa: Ala)
FT		/note= "This coding sequence contains one intron"
FT	intron	238. .3040
FT		/*tag= c
FT		/number= 1
FT	exon	3041. .4546
FT		/*tag= d
FT		/number= 2
FT		/note= "This sequence is also found in ABT43852"

XX
PN WO2003050311-A2.
XX

PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR P-PSDB; ABJ39148.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2c; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polynucleotide sequence is the partial sequence of the MHC region that
CC contains 2 exons coding for the human CDSN gene of the invention
XX
SQ Sequence 4740 BP; 1050 A; 1305 C; 1207 G; 1178 T; 0 U; 0 Other;

Query Match 94.5%; Score 1503.2; DB 9; Length 4740;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3039 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3098

Qy       143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3099 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3158

Qy       203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3159 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3218

Qy       263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3219 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3278

Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      3279 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3338

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3339 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3398

Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3399 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3458

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      3459 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3518

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3519 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3578

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3579 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3638

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3639 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3698

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3699 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3758

Qy       803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3759 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3818

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3819 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3878

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3879 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3938

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3939 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3998

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3999 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4058

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4059 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4118

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4119 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4178

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4179 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4238

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4239 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4298

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4299 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4358

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4359 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4418

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4419 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4478

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4479 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4538

Qy      1583 GTCCATAA 1590
             ||||||||
Db      4539 GTCCATAA 4546
```

```
RESULT 13
AAH45309
ID   AAH45309 standard; DNA; 5337 BP.
XX
AC   AAH45309;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human MHC S DNA.
XX
KW   Human; MHC S; major histocompatibility complex S; vulgar psoriasis;
KW   diagnosis; primer; SEEK1; HCR; a-helix coiled-coil rod homologue;
KW   polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1. .99
FT                   /*tag= a
FT   intron          100. .2903
FT                   /*tag= b
FT   exon            2904. .5337
FT                   /*tag= c
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Claim 1; Page 41-44; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is the human major
CC   histocompatibility complex (MHC) S gene, which was used to design primers
CC   for use in the invention
XX
SQ   Sequence 5337 BP; 1184 A; 1487 C; 1315 G; 1351 T; 0 U; 0 Other;

  Query Match             94.5%;  Score 1503.2;  DB 4;  Length 5337;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220
```

```
Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4401 GTCCATAA 4408
```

```
RESULT 14
AAH45310
ID   AAH45310 standard; DNA; 25235 BP.
XX
AC   AAH45310;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human SEEK1 DNA.
XX
KW   Human; MHC S; major histocompatibility complex S; vulgar psoriasis;
KW   diagnosis; primer; SEEK1; HCR; a-helix coiled-coil rod homologue;
KW   polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1. .420
FT                   /*tag= a
FT   intron          421. .1281
FT                   /*tag= b
FT   exon            1282. .1405
FT                   /*tag= c
FT   intron          1406. .1601
FT                   /*tag= d
FT   exon            1602. .1702
FT                   /*tag= e
FT   intron          1703. .2351
FT                   /*tag= f
FT   exon            2352. .2364
FT                   /*tag= g
FT   intron          2365. .6286
FT                   /*tag= h
FT   exon            6287. .6509
FT                   /*tag= i
FT   intron          6510. .10416
FT                   /*tag= j
FT   exon            10417. .10493
FT                   /*tag= k
FT   intron          10494. .14243
FT                   /*tag= l
FT   exon            14244. .14407
FT                   /*tag= m
FT   intron          14408. .14243
FT                   /*tag= n
FT   exon            25190. .25235
FT                   /*tag= o
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Claim 1; Page 46-62; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is the human SEEK1 gene
CC   which was used to design primers for use in the invention
XX
SQ   Sequence 25235 BP; 5967 A; 6703 C; 6487 G; 6078 T; 0 U; 0 Other;

  Query Match             94.5%;  Score 1503.2;  DB 4;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271
```

```
Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
              ||||||||
Db      24052 GTCCATAA 24059
```

RESULT 15
ABQ75680
ID ABQ75680 standard; DNA; 55050 BP.
XX
AC ABQ75680;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human SEEK1 consensus genomic DNA.
XX
KW Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW antipsoriatic; gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.
XX
PT Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT gene, for treating, diagnosing or determining susceptibility to SEEK1
PT mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS Claim 4; Fig 2; 69pp; English.
XX
CC The present invention describes a polynucleotide (I) comprising: (a) a
CC nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC acid sequence (IV) that hybridises under stringent conditions to the
CC polynucleotide of (II) or (III). (I) has antiinflammatory and
CC antipsoriatic activities and can be used in gene therapy. The alleles of
CC one or more polymorphisms in the SEEK1 gene can be used for the
CC manufacture of a medicament for the diagnosis and treatment of SEEK1
CC mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC prognostic or therapeutic methods, or as research tools in drug
CC screening. Human SEEK1 is located to chromosome 6p21. The present
CC sequence represents a human SEEK1 nucleotide sequence
XX
SQ Sequence 55050 BP; 13076 A; 14198 C; 14560 G; 13197 T; 0 U; 19 Other;

```
  Query Match             94.5%;  Score 1503.2;  DB 6;  Length 55050;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47898 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 47957

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47958 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 48017

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48018 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 48077

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48078 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 48137

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      48138 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 48197

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48198 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 48257

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48258 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 48317

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      48318 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 48377

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48378 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 48437
```

```
Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48438 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 48497

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48498 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 48557

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48558 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 48617

Qy       803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48618 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 48677

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48678 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 48737

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48738 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 48797

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48798 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 48857

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48858 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 48917

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48918 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 48977

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48978 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 49037

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCCACCAGGCA 1282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49038 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCCACCAGGCA 49097

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49098 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 49157

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49158 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 49217

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49218 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 49277

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49278 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 49337

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49338 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 49397

Qy      1583 GTCCATAA 1590
            ||||||||
Db     49398 GTCCATAA 49405
```

Search completed: March 21, 2006, 22:49:52
Job time : 1006 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:06:26 ; Search time 1369 Seconds
(without alignments)
9604.318 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1588.4	99.9	1590	9	US-10-677-865-2	Sequence 2, Appli
c 2	1586.8	99.8	1590	6	US-10-204-884-4	Sequence 4, Appli
3	1586.8	99.8	1725	6	US-10-316-175-14	Sequence 14, Appl
4	1580	99.4	1725	6	US-10-164-230-60	Sequence 60, Appl
5	1579.6	99.3	1590	6	US-10-204-884-2	Sequence 2, Appli
6	1570.8	98.8	2284	9	US-10-883-805-12	Sequence 12, Appl
7	1570	98.7	2530	3	US-09-968-007A-678	Sequence 678, App
8	1570	98.7	2530	6	US-10-316-175-11	Sequence 11, Appl
9	1570	98.7	2530	8	US-10-723-860-4323	Sequence 4323, Ap
10	1570	98.7	2530	9	US-10-843-641A-7148	Sequence 7148, Ap
11	1503.2	94.5	4740	6	US-10-316-175-13	Sequence 13, Appl
12	1503.2	94.5	5337	6	US-10-164-230-1	Sequence 1, Appli
13	1503.2	94.5	5700	6	US-10-204-884-1	Sequence 1, Appli
14	1503.2	94.5	25235	6	US-10-164-230-2	Sequence 2, Appli
15	1503.2	94.5	55050	7	US-10-467-752-4	Sequence 4, Appli
16	1501.2	94.4	1506	6	US-10-316-175-68	Sequence 68, Appl
17	1411.4	88.8	2551	8	US-10-723-860-8160	Sequence 8160, Ap
c 18	949.6	59.7	3673778	6	US-10-312-841-1	Sequence 1, Appli
19	745.4	46.9	3673778	6	US-10-312-841-2	Sequence 2, Appli
20	655.8	41.2	659	6	US-10-029-386-25367	Sequence 25367, A
21	505	31.8	505	6	US-10-029-386-11663	Sequence 11663, A
22	323.4	20.3	580	4	US-09-925-065A-244963	Sequence 244963,
23	323	20.3	580	4	US-09-925-065A-244955	Sequence 244955,
24	323	20.3	580	4	US-09-925-065A-244956	Sequence 244956,
25	85	5.3	99	6	US-10-316-175-67	Sequence 67, Appl
26	84.6	5.3	85	6	US-10-204-884-147	Sequence 147, App
27	84.2	5.3	85	6	US-10-204-884-146	Sequence 146, App
28	80.6	5.1	81	6	US-10-204-884-141	Sequence 141, App
29	80.6	5.1	81	6	US-10-204-884-150	Sequence 150, App
30	80.6	5.1	81	7	US-10-467-752-76	Sequence 76, Appl
31	79.6	5.0	80	6	US-10-204-884-144	Sequence 144, App
32	79	5.0	79	6	US-10-204-884-145	Sequence 145, App
33	78.6	4.9	79	6	US-10-204-884-143	Sequence 143, App
34	78.6	4.9	79	6	US-10-204-884-149	Sequence 149, App
35	74.8	4.7	76	7	US-10-467-752-73	Sequence 73, Appl
36	74.6	4.7	75	6	US-10-204-884-151	Sequence 151, App
37	74.6	4.7	75	7	US-10-467-752-74	Sequence 74, Appl
38	74	4.7	76	6	US-10-204-884-152	Sequence 152, App
39	73.6	4.6	74	6	US-10-204-884-156	Sequence 156, App
40	73.2	4.6	74	7	US-10-467-752-70	Sequence 70, Appl
41	71.6	4.5	72	6	US-10-204-884-148	Sequence 148, App
42	71.6	4.5	72	6	US-10-204-884-157	Sequence 157, App
43	71.2	4.5	72	7	US-10-467-752-69	Sequence 69, Appl
44	70.6	4.4	71	6	US-10-204-884-127	Sequence 127, App
45	70.6	4.4	71	7	US-10-467-752-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1
US-10-677-865-2
; Sequence 2, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/10/677,865
; CURRENT FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: US/09/381,656
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Human
US-10-677-865-2

Query Match 99.9%; Score 1588.4; DB 9; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
```

```
Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db       361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy      1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db      1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
Db      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
US-10-204-884-4/c
; Sequence 4, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-4

Query Match 99.8%; Score 1586.8; DB 6; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db      1590 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 1531

Qy        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db      1530 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 1471

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db      1470 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 1411

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db      1410 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 1351

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db      1350 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 1291

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db      1290 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 1231
```

```
Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db      1230 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 1171

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db      1170 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 1111

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db      1110 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 1051

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db      1050 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 991

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db       990 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 931

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db       930 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 871

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db       870 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 811

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db       810 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 751

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db       750 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 691

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db       690 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 631

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db       630 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 571

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db       570 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 511

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db       510 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 451

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db       450 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 391

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db       390 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 331

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       330 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 271

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       270 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 211

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       210 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 151

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       150 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 91

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db         90 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 31

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
Db         30 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1
```

RESULT 3
US-10-316-175-14
; Sequence 14, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-316-175-14

Query Match 99.8%; Score 1586.8; DB 6; Length 1725;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494
```

```
Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

```
              ||||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604


RESULT 4
US-10-164-230-60
; Sequence 60, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
;   LENGTH: 1725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (15)...(1601)
US-10-164-230-60

  Query Match             99.4%;  Score 1580;  DB 6;  Length 1725;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1577;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy           1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy          61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy         121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy         181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy         241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy         301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy         361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
               |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy         421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
               |||||||:|||||||||||||||||||||| ||||:|||||||||||||||||||||||||
Db         435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy         481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy         541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614
```

```
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              :|||||||||||||||||||:|||:||:||||||||||||||||||||||||||||||||
Db       1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||:||||||||||
Db       1575 CCCCAAGGAGAGTTACTCRACAGTCCATAA 1604
```

RESULT 5
US-10-204-884-2
; Sequence 2, Application US/10204884

; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1590)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (447)..(450)
; OTHER INFORMATION: Ser deletion
US-10-204-884-2

Query Match 99.3%; Score 1579.6; DB 6; Length 1590;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1570; Conservative 18; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGWTGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||:|||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        121 CCYTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCYTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGYTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAARCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAR 600
```

```
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCYGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGRCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAYCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCRTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              :|||||||||||||||||||:|||:||:||||||||||||||||||||||||||||||||
Db       1201 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCSAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGYGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||:||||||||||
Db       1561 CCCCAAGGAGAGTTACTCRACAGTCCATAA 1590


RESULT 6
US-10-883-805-12
; Sequence 12, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 12
;   LENGTH: 2284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-12

  Query Match             98.8%;  Score 1570.8;  DB 9;  Length 2284;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1578;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
```

```
Db        587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Db        647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Db        707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Db        767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Db        827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Db        887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Db        947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Db       1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Db       1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Db       1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Db       1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Db       1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Db       1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Db       1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Db       1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Db       1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Db       1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590

Db       1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636
```

RESULT 7

```
US-09-968-007A-678
; Sequence 678, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Usining Sign
;  TITLE OF INVENTION:  Gene Sets
;  FILE REFERENCE: 689290-71
;  CURRENT APPLICATION NUMBER: US/09/968,007A
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,172
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,173
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,278
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,294
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,295
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,316
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 1001
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 678
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-968-007A-678

  Query Match           98.7%;  Score 1570;  DB 3;  Length 2530;
  Best Local Similarity  99.6%;  Pred. No. 0;
  Matches 1584;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614
```

```
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

RESULT 8
US-10-316-175-11
; Sequence 11, Application US/10316175

```
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-316-175-11
```

Query Match 98.7%; Score 1570; DB 6; Length 2530;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734
```

```
Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            ||||||||||||||||||||||||||||||
Db      1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603


RESULT 9
US-10-723-860-4323
; Sequence 4323, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4323
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-4323

  Query Match             98.7%;  Score 1570;  DB 8;  Length 2530;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1584;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
```

```
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603


RESULT 10
US-10-843-641A-7148
; Sequence 7148, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7148
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2530)
;   OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-7148

  Query Match             98.7%;  Score 1570;  DB 9;  Length 2530;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1584;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC-G 1513

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1573

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1574 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1603
```

```
RESULT 11
US-10-316-175-13
; Sequence 13, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 4740
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-316-175-13

  Query Match             94.5%;  Score 1503.2;  DB 6;  Length 4740;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3039 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3098

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3099 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3158

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3159 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3218

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3219 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3278

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
               ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        3279 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3338

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3339 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3398

Qy         443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3399 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3458

Qy         503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
               |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        3459 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3518

Qy         563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3519 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3578

Qy         623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3579 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3638

Qy         683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3639 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3698

Qy         743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3699 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3758

Qy         803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3759 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3818

Qy         863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3819 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3878

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3879 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3938

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3939 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3998

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3999 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4058

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4059 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4118

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4119 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4178

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4179 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4238

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4239 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4298

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4299 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4358

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4359 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4418

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4419 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4478

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4479 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4538

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4539 GTCCATAA 4546


RESULT 12
US-10-164-230-1
; Sequence 1, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 5337
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(99)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (2904)...(5337)
US-10-164-230-1

  Query Match             94.5%;  Score 1503.2;  DB 6;  Length 5337;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy         443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy         503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy         563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy         623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy         683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy         743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy         803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3621 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy         863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy         923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy         983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
             ||||||||
Db       4401 GTCCATAA 4408
```

RESULT 13
US-10-204-884-1
; Sequence 1, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-1

Query Match 94.5%; Score 1503.2; DB 6; Length 5700;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3299 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3358

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3359 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3418

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3419 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3478

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3479 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3538

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3539 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3598

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3599 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3658

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3659 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3718

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3719 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3778

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3779 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3838

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3839 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3898

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3899 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3958

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3959 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 4018

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4019 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 4078

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4079 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 4138

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4139 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 4198

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 4258

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4318

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4378

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4379 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4438
```

```
Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4439 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4498

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4499 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4558

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4559 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4618

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4619 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4678

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4679 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4738

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4739 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4798

Qy       1583 GTCCATAA 1590
             ||||||||
Db       4799 GTCCATAA 4806


RESULT 14
US-10-164-230-2
; Sequence 2, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 25235
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(420)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1282)...(1405)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1702)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1631)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (2352)...(2364)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (6287)...(6509)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (10417)...(10493)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14407)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14344)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (25190)...(25235)
US-10-164-230-2

  Query Match             94.5%;  Score 1503.2;  DB 6;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451
```

```
Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
              ||||||||
Db      24052 GTCCATAA 24059
```

```
RESULT 15
US-10-467-752-4
; Sequence 4, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 55050
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13351)..(13351)
;   OTHER INFORMATION: n is a or t or g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13401)..(13401)
;   OTHER INFORMATION: n is a or t or g or c
US-10-467-752-4

  Query Match             94.5%;  Score 1503.2;  DB 7;  Length 55050;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47898 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 47957

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47958 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 48017

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48018 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 48077

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48078 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 48137

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      48138 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 48197

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48198 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 48257

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48258 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 48317

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      48318 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 48377

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48378 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 48437

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48438 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 48497

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48498 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 48557

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48558 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 48617

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48618 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 48677

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48678 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 48737

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48738 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 48797

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48798 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 48857

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      48858 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 48917

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48918 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 48977

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48978 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 49037

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49038 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 49097

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49098 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 49157

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49158 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 49217

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49218 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 49277

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49278 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 49337

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49338 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 49397

Qy       1583 GTCCATAA 1590
              ||||||||
Db      49398 GTCCATAA 49405


Search completed: March 22, 2006, 02:30:13
Job time : 1377 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:07:45 ; Search time 580 Seconds
(without alignments)
6392.090 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	323.4	20.3	580	6	US-09-925-065A-244963	Sequence 244963,
	2	323	20.3	580	6	US-09-925-065A-244955	Sequence 244955,
	3	323	20.3	580	6	US-09-925-065A-244956	Sequence 244956,
	4	60.8	3.8	1475	8	US-10-750-185-59559	Sequence 59559, A
	5	60.8	3.8	1475	8	US-10-750-623-59559	Sequence 59559, A
	6	57.2	3.6	2183	9	US-11-072-512-1064	Sequence 1064, Ap
	7	51.6	3.2	30191	7	US-10-330-773-631	Sequence 631, App
c	8	49	3.1	153376	12	US-11-121-086-5	Sequence 5, Appli
	9	47.2	3.0	204803	7	US-10-330-773-325	Sequence 325, App
	10	45.4	2.9	2479	12	US-11-136-527-3303	Sequence 3303, Ap
c	11	44.6	2.8	3240	12	US-11-052-554A-529	Sequence 529, App
	12	44.2	2.8	1097	9	US-11-096-568A-18153	Sequence 18153, A
	13	44.2	2.8	1375	9	US-11-096-568A-18992	Sequence 18992, A
	14	44	2.8	671	9	US-11-031-206-21	Sequence 21, Appl
	15	43.4	2.7	1734	7	US-10-063-703-51	Sequence 51, Appl
	16	43.4	2.7	1734	9	US-11-103-195-51	Sequence 51, Appl
	17	43.4	2.7	1734	12	US-11-102-240-51	Sequence 51, Appl
	18	43.4	2.7	72352	7	US-10-330-773-43	Sequence 43, Appl
	19	43	2.7	1707	7	US-10-932-182A-80360	Sequence 80360, A
	20	43	2.7	1707	7	US-10-932-182A-80360	Sequence 80360, A
c	21	43	2.7	47419	7	US-10-330-773-654	Sequence 654, App
	22	42.6	2.7	573	6	US-09-925-065A-236102	Sequence 236102,
	23	42	2.6	600	12	US-11-136-527-7399	Sequence 7399, Ap
c	24	41	2.6	833	8	US-10-750-185-56670	Sequence 56670, A
c	25	41	2.6	833	8	US-10-750-623-56670	Sequence 56670, A
	26	41	2.6	1634	9	US-11-096-568A-11426	Sequence 11426, A
c	27	41	2.6	3921	12	US-11-052-554A-523	Sequence 523, App
	28	40.8	2.6	2274	9	US-11-126-022-31	Sequence 31, Appl
	29	40.8	2.6	2404	9	US-11-072-512-1740	Sequence 1740, Ap
	30	40.8	2.6	3850	9	US-11-126-022-40	Sequence 40, Appl
c	31	40.4	2.5	2562	12	US-11-052-554A-533	Sequence 533, App
c	32	40	2.5	552	6	US-09-925-065A-373740	Sequence 373740,
	33	40	2.5	663	6	US-09-925-065A-3260	Sequence 3260, Ap
	34	39.8	2.5	1274	9	US-11-096-568A-21617	Sequence 21617, A
	35	39.8	2.5	5706	12	US-11-052-554A-519	Sequence 519, App
c	36	39	2.5	2765	9	US-11-072-512-528	Sequence 528, App
	37	39	2.5	20377	7	US-10-330-773-692	Sequence 692, App
	38	38.8	2.4	2623	8	US-10-750-185-63292	Sequence 63292, A
	39	38.8	2.4	2623	8	US-10-750-623-63292	Sequence 63292, A
	40	38.4	2.4	10361	7	US-10-330-773-629	Sequence 629, App
	41	38.4	2.4	127943	7	US-10-330-773-628	Sequence 628, App
c	42	38.2	2.4	1503	12	US-11-013-247A-3	Sequence 3, Appli
c	43	38.2	2.4	1587	12	US-11-013-247A-18	Sequence 18, Appl
c	44	38.2	2.4	1648	9	US-11-072-175-43	Sequence 43, Appl
	45	38	2.4	494	12	US-11-128-061-2640	Sequence 2640, Ap

ALIGNMENTS

RESULT 1
US-09-925-065A-244963
; Sequence 244963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244963
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-244963

Query Match 20.3%; Score 323.4; DB 6; Length 580;
Best Local Similarity 99.7%; Pred. No. 9.5e-74;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy       143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 375

Qy       203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 555

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAG 407
              |||||||||||||||||||||||||
Db        556 CATCCGGTTCCTCCCAGCTGGGGAG 580


RESULT 2
US-09-925-065A-244955
; Sequence 244955, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244955
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244955

  Query Match             20.3%;  Score 323;  DB 6;  Length 580;
  Best Local Similarity   99.4%;  Pred. No. 1.2e-73;
  Matches  323;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCYTGTAAGGACCCCACGCGTA 315

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 375

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 555

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAG 407
              |||||||||||||||||||||||||
Db        556 CATCCGGTTCCTCCCAGCTGGGGAG 580


RESULT 3
US-09-925-065A-244956
; Sequence 244956, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244956
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244956

  Query Match             20.3%;  Score 323;  DB 6;  Length 580;
  Best Local Similarity   99.4%;  Pred. No. 1.2e-73;
  Matches  323;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGYTTCAGTAGCT 375

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 555

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAG 407
              |||||||||||||||||||||||||
Db        556 CATCCGGTTCCTCCCAGCTGGGGAG 580


RESULT 4
US-10-750-185-59559
; Sequence 59559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
```

```
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 59559
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Bovine  19866881051537
US-10-750-185-59559

  Query Match             3.8%;  Score 60.8;  DB 8;  Length 1475;
  Best Local Similarity   80.7%;  Pred. No. 7.8e-06;
  Matches   71;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||| || || ||| ||  ||| ||| ||||||| |||| ||| || ||||  ||
Db        221 ATGGGCTCTTCACGAGCATCCCAGATAGGGTGTGTGGGAGGGCGAGGGGTGCTGGCTTTG 280

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGA 88
              |||||||||||||||||||||| ||| |
Db        281 CTGCTGGCTGGTCTCCTCCTGCAAGGTA 308


RESULT 5
US-10-750-623-59559
; Sequence 59559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 59559
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Bovine  19866881051537
US-10-750-623-59559

  Query Match             3.8%;  Score 60.8;  DB 8;  Length 1475;
  Best Local Similarity   80.7%;  Pred. No. 7.8e-06;
  Matches   71;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||| || || ||| ||  ||| ||| ||||||| |||| ||| || ||||  ||
Db        221 ATGGGCTCTTCACGAGCATCCCAGATAGGGTGTGTGGGAGGGCGAGGGGTGCTGGCTTTG 280

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGA 88
              |||||||||||||||||||||| ||| |
Db        281 CTGCTGGCTGGTCTCCTCCTGCAAGGTA 308


RESULT 6
US-11-072-512-1064
; Sequence 1064, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1064
;   LENGTH: 2183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-072-512-1064

  Query Match             3.6%;  Score 57.2;  DB 9;  Length 2183;
  Best Local Similarity   46.3%;  Pred. No. 7.3e-05;
  Matches  188;  Conservative    0;  Mismatches  218;  Indels    0;  Gaps    0;

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              |||   ||| |  |      |  | | |   ||  ||   | | | ||  |   | |  |
Db        564 AGCCACAGCAGCAGCAGCCCCAGCAGCAGCCACAGCAGCAGCAGCCCCAGCAGCAGCAGC 623

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
               ||  |  ||     |   |  ||| ||| ||  ||| |||  |||   |  ||||||
Db        624 TCCACCAGCAGCCCCAGTAGCAGCTCTAGCAGCAGCAGCAGCAGCAGCCCCAGCAGCAGC 683

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
                |||   ||  |||||   ||| ||| |  |  || |||||||  | |||  ||||
Db        684 AACTCCAGTAGCAGCTCCAGCAGCAGCAGCCCCAGCAGTAGCTCCAGCAGCAGCAGCAGC 743

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              |  |||| |    |  |  | |  |    |  |  |     | |  |     ||
Db        744 AGCCCCAGTAGCAGCTCCAGCAGCCCCAGTAGCAGCTCCAGCAGCAGCAGCAGCAGCCCC 803

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
              |||||  |  |  |     | ||   ||       |   ||| |  |  |
Db        804 AGTAGCAGCTCCAGCAGCCCCAGTAGCAGCTCTAGCAGCAGCAGCTCCAGCAGCAGCAGT 863

Qy        400 CTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 459
              |   | |||||||||| |         |  |||||  || |         | | |||||||  |
Db        864 CCCAGCAGCAGCAGCCCCAGTAGCAGCGGCAGCAGCCCCAGTAGCAGCAACAGCAGCCCC 923

Qy        460 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 505
                |   ||||||||||||  |  |    ||||||||||  |  ||
Db        924 AGTAGTAGCAGCAGCAGCCCCAGTAGCAGCAGCAGCAGCCCCAGCC 969


RESULT 7
US-10-330-773-631
; Sequence 631, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
```

```
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 631
;   LENGTH: 30191
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(30191)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-631

  Query Match             3.2%;  Score 51.6;  DB 7;  Length 30191;
  Best Local Similarity   45.9%;  Pred. No. 0.0039;
  Matches  177;  Conservative    0;  Mismatches  209;  Indels    0;  Gaps    0;

Qy       1008 CCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCAT 1067
                | | | || | |  |  |  || |  |||     ||   | || |||| ||    ||
Db      14115 CTCCTGCTCTTCCTCCTGCTCCTGCTCCTCCTCCTGCTCCTGCTCCTCCTGCTCTTCCTC 14174

Qy       1068 CATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGG 1127
              |  | || ||   |  |  | ||||  ||    | | |    |       |     |
Db      14175 CTGCTCCTGCTCCTGCTCCTCCTCCTGCTCCTGCTCCTCCTGCTCCTGCTCTTCCTCCTG 14234

Qy       1128 GGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTC 1187
                |     ||       ||  || | |  |||||     |  |||||||| |||| | ||
Db      14235 CTCCTGCTCCTCCTCCTGCTCCTGCTCCTGCTCCTCCTGCTCCTGCTCTTCCTCCTGCTC 14294

Qy       1188 TCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGG 1247
                |    ||  | | ||   |   |||  | || || | ||  ||| |    |||||
Db      14295 CTGCTCTTCCTCCTGCTCCTCCTGCTCCTCCTGCTCCTGCTCCTCCTCCTGCTCCTGCTC 14354

Qy       1248 CAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAG 1307
              | |  |||||    || ||||||||    |   |  |  || || |  |  | ||||| |
Db      14355 CTGCTCTTCCTCCTGCTCCTGCTCCTGCTCCTCCTGCTCCTGCTCCTCCTCCTGCTCCTG 14414

Qy       1308 TTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGG 1367
                |||   ||  |          | |        | ||   |  |   ||| |||| ||
Db      14415 CTCCTGCTCCTGCTCTTCCTCCTGCTCCTGCTCCTGCTCCTCCTGCTCCTGCTCCTCCTC 14474

Qy       1368 TCACCCTTGCATGTCTGTCTCCTCCT 1393
                | | |||   ||   || |||||
Db      14475 CTGCTCCTGCTCTTCCTCCTGCTCCT 14500


RESULT 8
US-11-121-086-5/c
; Sequence 5, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
;   LENGTH: 153376
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-5

  Query Match             3.1%;  Score 49;  DB 12;  Length 153376;
  Best Local Similarity   46.8%;  Pred. No. 0.027;
  Matches  154;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203
              | |  |||| |  ||||   |  ||| |||  ||||| ||||      | ||  |
Db      16368 GCCTCCTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCCTCC 16309

Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263
                | | |  |||  |  | || || |  ||| |    ||| |  | || | |||     |
Db      16308 TGTGCCTCCTCCTCCTCCTCCTGTGCCTCCTCCTCCTCCTCCTCCTGTTCCTCCTCCTCC 16249

Qy       1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323
              |||| ||| |  |   |  || | |||| |  |  |  ||      |||   ||     |
Db      16248 TCCTGTTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCCTGT 16189

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
                |  |       | |||| |  |   ||| ||| |||       | | | |   | |
Db      16188 GCTTCCTCCTCCTCCTCCTGTGCCTCCTCCTCCTGCTGTTCCTCCTCCTCCTCCTCCTGT 16129

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443
                |||||||  |     |  |  | ||     |||  ||  | |     ||||| |
Db      16128 TCCTCCTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCCTCCTGTTCCTCCTCCTCCTCC 16069

Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGC 1472
              || ||    |||    ||||  | || |
Db      16068 TGTTCCTCCTCCTCCTCCTCCTCCTGTTC 16040


RESULT 9
US-10-330-773-325
; Sequence 325, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 325
;   LENGTH: 204803
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(204803)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-325

  Query Match             3.0%;  Score 47.2;  DB 7;  Length 204803;
  Best Local Similarity   46.8%;  Pred. No. 0.085;
  Matches  148;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy        186 CAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAG 245
              |||| |   |||||||  |||  ||  ||||||     |||||  ||      ||    ||
Db      92075 CAGCAGTAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGCAGTAG 92134

Qy        246 AAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTC 305
              |||  | |  |  ||  |||  |||||  |||| |   |||||  |  |     |  ||
Db      92135 CAGCAGTAGCAGCAGCAGCAGCAGTAGCAGCAGCAGCAGCAGTAGCAGTAGCAGTAG 92194

Qy        306 TGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGG 365
                |       |      | |  |      |     |||| |  |  | | |  | |
Db      92195 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGCAGCAGCAGCAGCAGTAG 92254

Qy        366 CTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGG 425
              |  ||       |      | |  |  |       | |||||||||  |  |    |
Db      92255 CAGTAGCAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGCAGTAG 92314

Qy        426 AAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCA 485
              |||||| ||     |   | || ||||||  |     |||||||| ||||||
Db      92315 CAGCAGCAGCAGCAGCAGTAGCAGTAGCAGCAGTAGCAGTAGCAGCAGTAGCAGCAGCAG 92374

Qy        486 GTTCAGCAGCAGCAGC 501
                |||||||||||||
Db      92375 CAGCAGCAGCAGCAGC 92390
```

```
RESULT 10
US-11-136-527-3303
; Sequence 3303, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3303
;   LENGTH: 2479
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-3303

  Query Match             2.9%;  Score 45.4;  DB 12;  Length 2479;
  Best Local Similarity   49.0%;  Pred. No. 0.084;
  Matches  177;  Conservative    0;  Mismatches  181;  Indels    3;  Gaps    2;

Qy        144 CACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTA 203
              || |  |   | |  |  | ||| ||   | | |||   || |  | ||     | ||||
Db       1794 CAGCCACAACAGCAGCGGCAGCCGCAGCAGCAGGGGCAGCTGCGACAGCAGGGGCAGCTG 1853

Qy        204 CAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTC 263
              |||  ||  ||||||    ||||||  |     |||  ||| |||   ||  |  ||
Db       1854 CAGCAGCGGCAGCAGCAGGGGCAGCCGCGACAGCAGCAGCAGCAGCAGGGGCAGCCGCGA 1913

Qy        264 CAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTT-CTGCAGGATCTTTTAAGC 322
              |||  | |||  |||| |   ||||   | |    | ||   | |||||  |      ||
Db       1914 CAGCAGCAGCAGCAGCAGGGGCAGCCGCGGCAGCAGCGGCAGCAGCAGGGGCAGCCGAGG 1973

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||  |  ||       |   ||     |   | | | | | |   |  |   | ||  |
Db       1974 CATCAGAGGCAGCAGCAGGGGCAGCCGCAGCTGCAGCAGCAGCAGCAGCAGCAGCAG 2033

Qy        383 CATCCGGTTCCTC--CCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCA 440
              ||   |  || |  |     |  ||||||||    |   || |||||||||
Db       2034 CAGGGGCAACCGCGACAGCAGCGTCAGCAGCAGGGGCAGCCTCGGCAGCAGCGGCAGCAG 2093

Qy        441 CTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAG 500
              |  ||| ||| |  || |       ||| ||||  |||||      | ||||||||||||
Db       2094 CAGGGGCAGCCGAGGCACCAGCGGCAGCTGCAGGGGCAGCCGCAGCTGCAGCAGCAGCAG 2153

Qy        501 C 501
              |
Db       2154 C 2154


RESULT 11
US-11-052-554A-529/c
; Sequence 529, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 529
;   LENGTH: 3240
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-529

  Query Match             2.8%;  Score 44.6;  DB 12;  Length 3240;
  Best Local Similarity   41.7%;  Pred. No. 0.14;
  Matches  275;  Conservative    0;  Mismatches  384;  Indels    0;  Gaps    0;

Qy        836 CAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGG 895
              | |||    || || || | ||  || | || | ||   ||  | | | |   || | |
Db        764 CCGGCCGCGCCGGTTCCGCCATGCCCGCCGTCGCCAGCGCCGGTTGCGCCGGAGATGCCG 705

Qy        896 TGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCT 955
              ||  |  |   | ||  |    ||  | | |||  ||  |  ||  | |    |
Db        704 ACGGCACCAGCCGCGCCATCTCCGCCGTCGCCGCCATTGCCAGCGATACCGTTGGCGCCC 645

Qy        956 ATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCT 1015
                ||   | | | ||   |   |              ||  | |  | | |   |
Db        644 GGAGTGGCACCGGCGCCGCCCGCGCCCGCGTCTCCGCCCTTTCCGCCGCTGCCGCCCGCG 585

Qy       1016 TTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCA 1075
              || ||||  ||  |   |||  ||  |   || ||||| |  | ||  | | ||
Db        584 CCGCCGCTGTTGCCGGCGGTGCTGCTGGTGACGCCTGCTCCACCGTCGCCGCCTTTGCCG 525

Qy       1076 GCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGC 1135
              |   | ||     || |   |              | || ||    |  |  |||
Db        524 CCATTGCCGCCCGCCCCGCCAGCCCCGGCGGTGCCGCCGTTGGCGGCGCCGCCGTCACCG 465

Qy       1136 TCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCC 1195
              | ||   | |||| ||||      |||  ||  ||  |   | | || || |    |
Db        464 CCGGTTCCGCCGGCGCCACCACTACCACCGGTGCCGCCGGACGCGCCGGTGCCTACGGGG 405

Qy       1196 CCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTT 1255
              || |  || |  |||   |||    ||||| | |  ||      |||  || |   | |
Db        404 CCGGTTGTGCCGGCAGCGCCAATGCCGCCGTTGCCGCCCTTACCCCCATCGCCGCCGGTC 345

Qy       1256 CCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAAT 1315
              |||   |  || |   |  | | || ||||||      | ||  | | || |  |   |
Db        344 CCCGCCGTGCCAGGAGCGGAGCCGGTACCGGCCTTGCCGGCGTCGCCGCCGTTGCCGCCT 285

Qy       1316 CGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTT 1375
              |  | |  | || |  |    | || |   |||    | | |  ||    | |
Db        284 TTACCGCCGGCACCACCGCTGCTGCCTGCGAGTGCACCGTCGCCGCCGGTGCCGCCTTTG 225

Qy       1376 GCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATC 1435
              |   || | ||      |  | | | |  ||     |  || | | |  | |
Db        224 CCGCCGGTGCCGCCGTTGCCGCCGTTGCCGCCGTTGCCTTGTTGGCCGCCGGTGCCGGTG 165

Qy       1436 CCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494
              ||  ||     |||    ||  || |||  | |  | |        | || ||| | ||
Db        164 CCGGCGGCTCCGCCGGTTCCGCCGGTTCCGCCGGCCCCGCCGGCCCCGCCGGCCCCGCC 106


RESULT 12
US-11-096-568A-18153
; Sequence 18153, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18153
;   LENGTH: 1097
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1097)
```

```
;   OTHER INFORMATION: Ceres Seq. ID no. 12363241
US-11-096-568A-18153

  Query Match             2.8%;  Score 44.2;  DB 9;  Length 1097;
  Best Local Similarity   49.8%;  Pred. No. 0.14;
  Matches  112;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy       1082 CGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTG 1141
              | ||||   |  ||  |   | || | || | |  | | | |    ||    ||     |
Db        128 CAGCAGACGCGGCGTTCTGCGTGTGCAAGACGGGCGTGCCCGACCAGGCGATGCAGGCGG 187

Qy       1142 GAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCA 1201
                || || | |  || |||||||  ||  ||   |||||   ||    | | |   |
Db        188 CGATCGACTACGCCTGCGCCAAGGGGGCCGACTGCGCCTCCGCCTCCAAGGGCGCGCCCT 247

Qy       1202 GTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA 1261
              | |   |||    || ||  |  |        || | |   ||||  |||    | |||||
Db        248 GCTACGGCAACGGCAACAAGGTGGCCGTCTGCTCCTACATCTGCAACAGCTACTACCAGA 307

Qy       1262 GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCA 1306
              |||  | || || |    | |||| ||  ||||| | ||  | |||
Db        308 GCCGCAGCGCCACGGGCGCCACCTGCGACTTCAACGGCGTCGCCA 352


RESULT 13
US-11-096-568A-18992
; Sequence 18992, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18992
;   LENGTH: 1375
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1375)
;   OTHER INFORMATION: Ceres Seq. ID no. 12368133
US-11-096-568A-18992

  Query Match             2.8%;  Score 44.2;  DB 9;  Length 1375;
  Best Local Similarity   49.8%;  Pred. No. 0.15;
  Matches  112;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy       1082 CGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTG 1141
              | ||||   |  ||  |   | || | || | |  | | | |    ||    ||     |
Db        409 CAGCAGACGCGGCGTTCTGCGTGTGCAAGACGGGCGTGCCCGACCAGGCGATGCAGGCGG 468

Qy       1142 GAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCA 1201
                || || | |  || |||||||  ||  ||   |||||   ||    | | |   |
Db        469 CGATCGACTACGCCTGCGCCAAGGGGGCCGACTGCGCCTCCGCCTCCAAGGGCGCGCCCT 528

Qy       1202 GTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA 1261
              | |   |||    || ||  |  |        || | |   ||||  |||    | |||||
Db        529 GCTACGGCAACGGCAACAAGGTGGCCGTCTGCTCCTACATCTGCAACAGCTACTACCAGA 588

Qy       1262 GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCA 1306
              |||  | || || |    | |||| ||  ||||| | ||  | |||
Db        589 GCCGCAGCGCCACGGGCGCCACCTGCGACTTCAACGGCGTCGCCA 633


RESULT 14
US-11-031-206-21
; Sequence 21, Application US/11031206
; Publication No. US20060031959A1
; GENERAL INFORMATION:
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Miao, Guo-Hua
;  APPLICANT: Falco, Saverio Carl
;  APPLICANT: Sakai, Hajime
;  APPLICANT: Famodu, Omolayo O.
;  APPLICANT: Odell, Joan T.
;  APPLICANT: Meyers, Blake
;  APPLICANT: Thorpe, Catherine
;  APPLICANT: Weng, Zude
;  TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
;  TITLE OF INVENTION:  Stress Response
;  FILE REFERENCE: BB1357 US NA
;  CURRENT APPLICATION NUMBER: US/11/031,206
;  CURRENT FILING DATE:  2005-01-07
;  PRIOR APPLICATION NUMBER: US/09/566,394
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: 60/133038
;  PRIOR FILING DATE: 1999-05-07
;  PRIOR APPLICATION NUMBER: 60/133042
;  PRIOR FILING DATE: 1999-05-07
;  PRIOR APPLICATION NUMBER: 60/133427
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/133437
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/133428
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/133438
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/133436
;  PRIOR FILING DATE: 1999-05-11
;  PRIOR APPLICATION NUMBER: 60/137667
;  PRIOR FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
;   LENGTH: 671
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-11-031-206-21

  Query Match             2.8%;  Score 44;  DB 9;  Length 671;
  Best Local Similarity   50.5%;  Pred. No. 0.14;
  Matches  107;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy       1137 CTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCC 1196
              |||  || || ||| |      || ||    |   ||||||   |||| | ||  |   |
Db        247 CTGCAGAAGCCGCTTCGGATTCTGCAGCTGCAGCTGCTCCTGCTCCTGCTCCGGCAACAG 306

Qy       1197 CAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTC 1256
              ||| || |  |||      | | |||  |    | | | |||| | | || | ||  | |
Db        307 CAGAAGCTGAAGCTGAAGCTGAAGCTGCTCCACCCACCACCACCGCGGCAGAAGCACCTG 366

Qy       1257 CCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 1316
              |||  ||  ||||| |     |    |  | || || |  ||| |||  |     ||||
Db        367 CCATCGCTGCTGCTGCTGCTCCTCCTGCTCCTGCTGCTCCAGAAGCAGCTCCGGCCAAGG 426

Qy       1317 GAGTGGCAAAATCATCCTTCAGCCTTGTGGCA 1348
                |  ||   || |||   ||   |  | |||
Db        427 CTGATGCTTGATGATCACACATAATGCTTGCA 458


RESULT 15
US-10-063-703-51
; Sequence 51, Application US/10063703
; Publication No. US20060008901A1
; GENERAL INFORMATION:
;  APPLICANT: Eaton,Dan L.
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Watanabe,Colin K.
```

```
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3230R1C1
;  CURRENT APPLICATION NUMBER: US/10/063,703
;  CURRENT FILING DATE:  2002-05-08
;      Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 51
;   LENGTH: 1734
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-063-703-51

  Query Match             2.7%;  Score 43.4;  DB 7;  Length 1734;
  Best Local Similarity   57.9%;  Pred. No. 0.25;
  Matches   77;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy        162 CTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTC 221
              | ||  || ||| |   |      |||| ||    |  ||  |||||||  |||||||
Db        996 CAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGGCAGCAGTGG 1055

Qy        222 TGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGG 281
              |||||||       |||||   | |||   ||||   ||  |||||    |||| | ||
Db       1056 TGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGG 1115

Qy        282 ATCCAGCATTGCC 294
              ||||||||  | |
Db       1116 ATCCAGCACCGGC 1128


Search completed: March 22, 2006, 02:39:56
Job time : 583 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 01:05:34 ; Search time 307 Seconds
(without alignments)
9206.264 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1588.4	99.9	1590	3	US-09-381-656-2	Sequence 2, Appli
	2	1580	99.4	1725	3	US-10-164-230-60	Sequence 60, Appl
	3	1503.2	94.5	5337	3	US-10-164-230-1	Sequence 1, Appli
	4	1503.2	94.5	25235	3	US-10-164-230-2	Sequence 2, Appli
	5	148.8	9.4	222	3	US-09-641-931F-35	Sequence 35, Appl
	6	58.8	3.7	7218	2	US-08-232-463-14	Sequence 14, Appl
	7	57.2	3.6	2183	3	US-10-104-047-1064	Sequence 1064, Ap
c	8	56.2	3.5	2793	2	US-08-209-747-1	Sequence 1, Appli
c	9	56.2	3.5	2793	2	US-08-458-298-1	Sequence 1, Appli
c	10	49.4	3.1	1926	3	US-09-249-585A-2	Sequence 2, Appli
c	11	49.4	3.1	1926	3	US-09-410-399-3	Sequence 3, Appli
c	12	49.4	3.1	2580	3	US-09-050-863-2	Sequence 2, Appli
c	13	49.4	3.1	2580	3	US-09-359-081-2	Sequence 2, Appli
	14	49.4	3.1	5452	2	US-09-130-114-1	Sequence 1, Appli
	15	49.4	3.1	8705	3	US-09-647-344A-14	Sequence 14, Appl
c	16	49.4	3.1	9600	3	US-08-910-647-1	Sequence 1, Appli
c	17	49.4	3.1	9600	3	US-09-620-925-1	Sequence 1, Appli
c	18	49.4	3.1	10596	2	US-07-884-811-15	Sequence 15, Appl
c	19	49.4	3.1	10596	2	US-07-885-971-15	Sequence 15, Appl
c	20	49.4	3.1	10596	2	US-08-087-783A-15	Sequence 15, Appl
c	21	49.4	3.1	10596	2	US-08-194-088B-15	Sequence 15, Appl
c	22	49.4	3.1	10596	2	US-08-194-087-15	Sequence 15, Appl
c	23	49.4	3.1	10596	6	PCT-US93-04648-15	Sequence 15, Appl
	24	49.4	3.1	16080	3	US-09-724-566A-48	Sequence 48, Appl
	25	49.4	3.1	16080	3	US-09-471-669A-48	Sequence 48, Appl
c	26	47	3.0	792	3	US-09-902-540-8438	Sequence 8438, Ap
c	27	47	3.0	8917	3	US-09-902-540-887	Sequence 887, App
c	28	46.2	2.9	786	2	US-08-403-379A-2	Sequence 2, Appli
c	29	46.2	2.9	786	2	US-08-929-414-2	Sequence 2, Appli
c	30	46	2.9	4180	3	US-09-814-915A-93	Sequence 93, Appl
c	31	46	2.9	4180	3	US-09-949-002-103	Sequence 103, App
	32	44.6	2.8	477	3	US-09-135-994-1	Sequence 1, Appli
	33	44.6	2.8	477	3	US-09-684-843A-1	Sequence 1, Appli
c	34	44.6	2.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	35	44.4	2.8	4174	3	US-09-949-002-138	Sequence 138, App
c	36	44.4	2.8	15044	3	US-09-949-002-675	Sequence 675, App
c	37	44.4	2.8	15044	3	US-09-949-002-710	Sequence 710, App
c	38	44.2	2.8	1235	2	US-08-557-309B-53	Sequence 53, Appl
c	39	44.2	2.8	2144	3	US-08-834-306-15	Sequence 15, Appl
c	40	44.2	2.8	2144	3	US-08-993-674A-15	Sequence 15, Appl
c	41	44.2	2.8	2144	3	US-09-256-976-15	Sequence 15, Appl
c	42	44	2.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	43	43.4	2.7	1734	3	US-10-012-231A-51	Sequence 51, Appl
	44	43.4	2.7	1734	3	US-10-015-389A-51	Sequence 51, Appl
	45	43.4	2.7	1734	3	US-10-006-768A-51	Sequence 51, Appl

ALIGNMENTS

Appl.

RESULT 1
US-09-381-656-2
; Sequence 2, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Human
US-09-381-656-2

Query Match 99.9%; Score 1588.4; DB 3; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1589; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
US-10-164-230-60
; Sequence 60, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)...(1601)
US-10-164-230-60

Query Match 99.4%; Score 1580; DB 3; Length 1725;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1577; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74
Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134
Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194
Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254
Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314
Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374
Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
```

```
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||:|||||||||||||||||||||| ||||:||||||||||||||||||||||||
Db        435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              :|||||||||||||||||||:|||:||:||||||||||||||||||||||||||||||||
Db       1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||:||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCRACAGTCCATAA 1604
```

RESULT 3
US-10-164-230-1
; Sequence 1, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)...(99)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2904)...(5337)
US-10-164-230-1

Query Match 94.5%; Score 1503.2; DB 3; Length 5337;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
               |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy         443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy         503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
```

```
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
              ||||||||

Db       4401 GTCCATAA 4408
```

```
RESULT 4
US-10-164-230-2
; Sequence 2, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 25235
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(420)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1282)...(1405)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1702)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1631)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (2352)...(2364)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (6287)...(6509)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (10417)...(10493)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14407)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14344)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (25190)...(25235)
US-10-164-230-2

  Query Match             94.5%;  Score 1503.2;  DB 3;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
                ||||||||
Db      24052 GTCCATAA 24059
```

RESULT 5
US-09-641-931F-35
; Sequence 35, Application US/09641931F
; Patent No. 6830885
; GENERAL INFORMATION:
; APPLICANT: Lanctot, et al.
; TITLE OF INVENTION: Nucleic Acid Molecule, Method and Kit for Selecting a
; TITLE OF INVENTION: Nucleic Acid Having A Desired Feature
; FILE REFERENCE: 2003390-0001
; CURRENT APPLICATION NUMBER: US/09/641,931F
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 222
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-641-931F-35

Query Match 9.4%; Score 148.8; DB 3; Length 222;
Best Local Similarity 82.2%; Pred. No. 3.1e-29;
Matches 171; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||| || || | ||||||| |||||||| ||||||| ||||||||| |||||||| ||
Db         15 ATGGGTTCCTCGCAGGCACCCCGGATGGGGAGTGTGGGAGGGCACGGGCTGATGGCATTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||| |||| |||||  | |||||||| | |||||||||||||||||| ||| |||| |||
Db         75 CTGATGGCCGGTCTTATTCTGCCAGGAATCTTGGCTAAGAGCATTGGGACCCTCTCGGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              |||||||||||||||| | |||||||||||| || ||||| || |||| ||| |||  |
Db        135 CCCTGTAAGGACCCCACGAGGATCACCTCCCCGAATGACCCTTGTCTCATTGGAAAGACT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTG 208
              | ||||| | || |||| ||| |  |||
Db        195 GGCTCCAACAGCATCAGCAGCCAAGGTG 222
```

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria

```
;       STATE:  VA
;       COUNTRY:  USA
;       ZIP:  22313-0299
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/232,463
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/935,313
;       FILING DATE:
;       APPLICATION NUMBER:  EP 91 114 300.6
;       FILING DATE:  26-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7218 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.7%;  Score 58.8;  DB 2;  Length 7218;
  Best Local Similarity   7.0%;  Pred. No. 5.2e-05;
  Matches   30;  Conservative  222;  Mismatches  174;  Indels    0;  Gaps    0;

Qy       1088 CTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCG 1147
              |||| |  |||||    |  |||             | | || | ||| ||   ||||
Db       1002 CTTCGTTTGCCATACGCTCACAGAATTAATTCCGAGCTTGGCTGCAGGTCGAGGGAGCTT 1061

Qy       1148 GCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTA 1207
              ||    :  ::::      :::: :::::::::  :::::   :::::  :  ::::
Db       1062 GCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1121

Qy       1208 GCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCT 1267
               : :::::  :  : ::  ::: :::: :: ::::: :  :  : ::::::    :::::
Db       1122 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1181

Qy       1268 GCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAA 1327
              :::::: ::   : ::  :::::::  :  :  ::::  :::::  ::   :  :
Db       1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241

Qy       1328 TCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCT 1387
              :: ::::::  :::: :  :  :   ::: :::::::  :: ::::: : : ::: :::
Db       1242 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1301

Qy       1388 CCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTG 1447
              :::::::  : ::  ::     ::::  :  :::::::: ::::  ::::::: ::  :
Db       1302 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1361

Qy       1448 CCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCC 1507
              ::   :::: :  ::::  :  : ::        :::::: :: :::: :::    : :::
Db       1362 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1421

Qy       1508 TAGCCC 1513
              : :::
Db       1422 YYYYYY 1427


RESULT 7
```

```
US-10-104-047-1064
; Sequence 1064, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. 6943241el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1064
;   LENGTH: 2183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-104-047-1064

  Query Match             3.6%;  Score 57.2;  DB 3;  Length 2183;
  Best Local Similarity   46.3%;  Pred. No. 8.1e-05;
  Matches  188;  Conservative    0;  Mismatches  218;  Indels    0;  Gaps    0;

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              |||    ||| |  |    |  | | |   ||  ||   | | | ||  |  | | |
Db        564 AGCCACAGCAGCAGCAGCCCCAGCAGCAGCCACAGCAGCAGCAGCCCCAGCAGCAGCAGC 623

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
              ||  |  ||      |   |   ||| ||| ||  ||| |||  |||   |  ||||||
Db        624 TCCACCAGCAGCCCCAGTAGCAGCTCTAGCAGCAGCAGCAGCAGCAGCCCCAGCAGCAGC 683

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
                |||   ||  |||||   ||| ||| |  |  || |||||||  | |||  ||||
Db        684 AACTCCAGTAGCAGCTCCAGCAGCAGCAGCCCCAGCAGTAGCTCCAGCAGCAGCAGCAGC 743

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              |  |||| |   |  |  |  |  |   |  | |    |  | |  |    ||
Db        744 AGCCCCAGTAGCAGCTCCAGCAGCCCCAGTAGCAGCTCCAGCAGCAGCAGCAGCCCC 803

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
                |||| |  |  |   | ||   ||       |   ||| |   |  |
Db        804 AGTAGCAGCTCCAGCAGCCCCAGTAGCAGCTCTAGCAGCAGCAGCTCCAGCAGCAGCAGT 863

Qy        400 CTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 459
              |   | |||||||| |         |  |||| || |        | | |||||| |
Db        864 CCCAGCAGCAGCAGCCCCAGTAGCAGCGGCAGCAGCCCCAGTAGCAGCAACAGCAGCCCC 923

Qy        460 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 505
              |   ||||||||||||  |  |    |||||||||  |  ||
Db        924 AGTAGTAGCAGCAGCAGCCCCAGTAGCAGCAGCAGCAGCCCCAGCC 969


RESULT 8
US-08-209-747-1/c
; Sequence 1, Application US/08209747
; Patent No. 5733771
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Colgin, Mark
;    TITLE OF INVENTION:  cDNAs Encoding Minor Ampullate Spider
;    TITLE OF INVENTION:  Silk Proteins
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/209,747
;       FILING DATE:  14-MAR-1994
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Murphy Jr., Gerald M.
;       REGISTRATION NUMBER:  28,977
;       REFERENCE/DOCKET NUMBER:  1447-104P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-205-8000
;       TELEFAX:  703-205-8050
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2793 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Nephila clavipes
;       TISSUE TYPE:  minor ampullate gland
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  183..2675
;       OTHER INFORMATION:  /product= "N. clavipes minor
;       OTHER INFORMATION:  ampullate silk protein"
US-08-209-747-1

  Query Match             3.5%;  Score 56.2;  DB 2;  Length 2793;
  Best Local Similarity   46.2%;  Pred. No. 0.00016;
  Matches  187;  Conservative    0;  Mismatches  218;  Indels    0;  Gaps    0;

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              ||||   ||| |  |||   | |||| |  | | |||     | ||||||  |      |
Db       2128 AGCACCTGCAGCGGCTCCAGCTCCTGCAGCGGCTCCAGCACCAGCACCTCTACCGTAACC 2069

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
               || ||  | | ||     |   |  |||| ||  ||| | |  |    ||||| |||
Db       2068 ACCAGCGCCTCCGGATCCTGCACCAGCAGCTGCAGCAGCACCAGCTCCTGCTCCGGCACC 2009

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
                || ||  |   ||  |||   |||||  |     || || |  | ||||||||
Db       2008 GTATCCACCTTGACCACCGTAGCCTCCAGCTCCTGCTCCTGCACCTGCAGCGGCTCCAGC 1949

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              || ||  | |   |     | |     || | |  |   | |       |||
Db       1948 TCCAGCACCAGCGGCTCCTGCACGACCACCTCTACTGTAACCCGCAGCGCCTCCAGCTCC 1889

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
                ||||  || |  | |  |  |||||             | ||||       ||
Db       1888 TGCTCCAGCAGCTGCCGCAGCACCAGCTCTTGCTCCGGCACCGTATCCACCCTGACCACC 1829

Qy        400 CTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 459
               | |   ||||  ||    ||      ||||||||||   |||  | | | |||||
Db       1828 GTAGCCTCCAGCTCCTGCTCCTGCACCTGCAGCGGCTCCAGCTCCAGCACCTGCAGCAGC 1769

Qy        460 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
                |||||  ||||| || |    | | |||||  |  || |
Db       1768 TCCAGCTCCAGCACCAGCACCTCTACCGTAACCAGCACCTCCTGC 1724


RESULT 9
US-08-458-298-1/c
; Sequence 1, Application US/08458298
; Patent No. 5756677
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Colgin, Mark
;    TITLE OF INVENTION:  cDNAs Encoding Minor Ampullate Spider
;    TITLE OF INVENTION:  Silk Proteins
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,298
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,747
;      FILING DATE:  14-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-104P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Nephila clavipes
;      TISSUE TYPE:  minor ampullate gland
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  183..2675
;      OTHER INFORMATION:  /product= "N. clavipes minor
;      OTHER INFORMATION:  ampullate silk protein"
US-08-458-298-1

  Query Match             3.5%;  Score 56.2;  DB 2;  Length 2793;
  Best Local Similarity   46.2%;  Pred. No. 0.00016;
  Matches  187;  Conservative    0;  Mismatches  218;  Indels    0;  Gaps    0;

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              ||||   ||| |  |||   | |||| |  | | |||     | ||||||  |      |
Db       2128 AGCACCTGCAGCGGCTCCAGCTCCTGCAGCGGCTCCAGCACCAGCACCTCTACCGTAACC 2069

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
               || ||  | | ||     |   |  |||| ||  ||| | |  |    ||||| |||
Db       2068 ACCAGCGCCTCCGGATCCTGCACCAGCAGCTGCAGCAGCACCAGCTCCTGCTCCGGCACC 2009

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
                || ||  |   ||  |||   |||||  |     || || |  | ||||||||
Db       2008 GTATCCACCTTGACCACCGTAGCCTCCAGCTCCTGCTCCTGCACCTGCAGCGGCTCCAGC 1949

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              || ||  | |   |     | |     || | |  |   | |       |||
Db       1948 TCCAGCACCAGCGGCTCCTGCACGACCACCTCTACTGTAACCCGCAGCGCCTCCAGCTCC 1889

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
                ||||  || |  | |  |  |||||             | ||||       ||
Db       1888 TGCTCCAGCAGCTGCCGCAGCACCAGCTCTTGCTCCGGCACCGTATCCACCCTGACCACC 1829

Qy        400 CTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 459
               | |   ||||  ||    ||      ||||||||||   |||  | | | |||||
Db       1828 GTAGCCTCCAGCTCCTGCTCCTGCACCTGCAGCGGCTCCAGCTCCAGCACCTGCAGCAGC 1769

Qy        460 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
```

```
                    ||||| ||||| || |    | | ||||| |  || |
Db       1768 TCCAGCTCCAGCACCAGCACCTCTACCGTAACCAGCACCTCCTGC 1724


RESULT 10
US-09-249-585A-2/c
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
;  APPLICANT: Horlick, Robert
;  TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
;  FILE REFERENCE: 0867/0D905
;  CURRENT APPLICATION NUMBER: US/09/249,585A
;  CURRENT FILING DATE:  1999-02-11
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1926
;   TYPE: DNA
;   ORGANISM: Epstein Barr Virus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1926)
;   OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

  Query Match             3.1%;  Score 49.4;  DB 3;  Length 1926;
  Best Local Similarity   45.4%;  Pred. No. 0.0085;
  Matches  222;  Conservative    0;  Mismatches  261;  Indels    6;  Gaps    1;

Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
                | ||||    |  | |   || |   |||   | |      || | ||  |  | ||  |
Db       1022 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 963

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
                | || ||   |   | | ||||   ||    | | |   ||      |  | |
Db        962 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 903

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
                |   | ||   |   | || || |  || ||    | ||||| |  ||| |   ||
Db        902 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 843

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
                | || ||   ||  ||   |||  | || || | |   ||| ||   |||||  |
Db        842 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 783

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
                | | |||     | ||||||||   ||   |  | |||| |  | ||  |||| | |
Db        782 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 723

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
                ||     |     ||     |  |       | ||   | ||   ||| || | ||  | |
Db        722 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 669

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
                | | | |   ||   | ||||||   |      |     |||||    || |  ||
Db        668 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 609

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
                |    |||||| |  |  ||   ||||   || ||  | |  |           | ||||||
Db        608 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 549

Qy       1490 GCTCCATCC 1498
                ||||  ||
Db        548 CCTCCTGCC 540


RESULT 11
US-09-410-399-3/c
; Sequence 3, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
;  APPLICANT:  Robertson, Erle S.
;  APPLICANT:  Cotter, Murray A.
;  TITLE OF INVENTION:  Methods to Inhibit or Enhance the Binding of Viral DNA
;  TITLE OF INVENTION:  to Genomic Host DNA
;  FILE REFERENCE:  UM-03778
;  CURRENT APPLICATION NUMBER: US/09/410,399
;  CURRENT FILING DATE:  1999-10-01
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1926
;   TYPE: DNA
;   ORGANISM: Epstein-Barr virus
US-09-410-399-3

  Query Match             3.1%;  Score 49.4;  DB 3;  Length 1926;
  Best Local Similarity   45.4%;  Pred. No. 0.0085;
  Matches  222;  Conservative    0;  Mismatches  261;  Indels    6;  Gaps    1;

Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
                | ||||    |  | |   || |   |||   | |      || | ||  |  | ||  |
Db       1022 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 963

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
                | || ||   |   | | ||||   ||    | | |   ||      |  | |
Db        962 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 903

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
                |   | ||   |   | || || |  || ||    | ||||| |  ||| |   ||
Db        902 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 843

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
                | || ||   ||  ||   |||  | || || | |   ||| ||   |||||  |
Db        842 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 783

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
                | | |||     | ||||||||   ||   |  | |||| |  | ||  |||| | |
Db        782 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 723

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
                ||     |     ||     |  |       | ||   | ||   ||| || | ||  | |
Db        722 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 669

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
                | | | |   ||   | ||||||   |      |     |||||    || |  ||
Db        668 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 609

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
                |    |||||| |  |  ||   ||||   || ||  | |  |           | ||||||
Db        608 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 549

Qy       1490 GCTCCATCC 1498
                ||||  ||
Db        548 CCTCCTGCC 540


RESULT 12
US-09-050-863-2/c
; Sequence 2, Application US/09050863
; Patent No. 6114111
;  GENERAL INFORMATION:
;    APPLICANT:  Lao, Ying
;    APPLICANT:  Hiang, Betty
;    APPLICANT:  Payan, Don
;    TITLE OF INVENTION:  Mammalian Protein Interaction Cloning
;    TITLE OF INVENTION:  System
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-4187
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,863
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Silva, Robin M.
;      REGISTRATION NUMBER:  38,304
;      REFERENCE/DOCKET NUMBER:  A-65638/DJB/RMS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 949-8711
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2580 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA
US-09-050-863-2

  Query Match             3.1%;  Score 49.4;  DB 3;  Length 2580;
  Best Local Similarity   45.4%;  Pred. No. 0.0096;
  Matches  222;  Conservative    0;  Mismatches  261;  Indels    6;  Gaps    1;

Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
             | ||||   |  | |   || |   |||   |  |      || | ||  |  | ||  |
Db       1405 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 1346

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
              | || ||   |   | | ||||   ||    | | |  ||      |  |  |
Db       1345 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 1286

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
              |   | ||   |   | || || |  || ||    |  ||||| |  ||| |   ||
Db       1285 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 1226

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
               | || ||   ||  ||   ||| | || || |  |  ||| ||   |||||  |
Db       1225 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 1166

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
              | | |||     | |||||||  ||  |  |  ||||| |  | ||  |||| | |
Db       1165 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 1106

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
              ||     ||    |  |          | ||   | ||   ||| || | ||  | ||
Db       1105 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 1052

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
              | | | |   ||  | ||||||   |      |     |||||   || |  ||
Db       1051 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 992

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
              |     |||||| |  |  ||   ||||  || ||  | |  |          | ||||||
Db        991 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 932

Qy       1490 GCTCCATCC 1498
              ||||  ||
Db        931 CCTCCTGCC 923


RESULT 13
US-09-359-081-2/c
; Sequence 2, Application US/09359081
; Patent No. 6316223
;   GENERAL INFORMATION:
;       APPLICANT: Lao, Ying
;                  Hiang, Betty
;                  Payan, Don
;       TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;                           System
;       NUMBER OF SEQUENCES: 5
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;            STREET: 4 Embarcadero Center, Suite 3400
;            CITY: San Francisco
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94111-4187
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/359,081
;            FILING DATE: 22-Jul-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/050,863
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Silva, Robin M.
;            REGISTRATION NUMBER: 38,304
;            REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 781-1989
;            TELEFAX: (415) 949-8711
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2580 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: unknown
;            TOPOLOGY: unknown
;       MOLECULE TYPE: DNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

  Query Match             3.1%;  Score 49.4;  DB 3;  Length 2580;
  Best Local Similarity   45.4%;  Pred. No. 0.0096;
  Matches  222;  Conservative    0;  Mismatches  261;  Indels    6;  Gaps    1;

Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
             | ||||   |  | |   || |   |||   |  |      || | ||  |  | ||  |
Db       1405 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 1346

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
              | || ||   |   | | ||||   ||    | | |  ||      |  |  |
Db       1345 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 1286

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
              |   | ||   |   | || || |  || ||    |  ||||| |  ||| |   ||
Db       1285 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 1226

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
               | || ||   ||  ||   ||| | || || |  |  ||| ||   |||||  |
Db       1225 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 1166

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
              | | |||     | |||||||  ||  |  |  ||||| |  | ||  |||| | |
Db       1165 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 1106

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
              ||     ||    |  |          | ||   | ||   ||| || | ||  | ||
Db       1105 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 1052

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
              | | | |   ||  | ||||||   |      |     |||||   || |  ||
Db       1051 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 992

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
```

```
Db        991 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 932

Qy       1490 GCTCCATCC 1498
Db        931 CCTCCTGCC 923
```

RESULT 14
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1

Query Match 3.1%; Score 49.4; DB 2; Length 5452;
Best Local Similarity 45.4%; Pred. No. 0.013;
Matches 222; Conservative 0; Mismatches 261; Indels 6; Gaps 1;

```
Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
Db       1400 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 1459

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
Db       1460 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 1519

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
Db       1520 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 1579

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
Db       1580 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 1639

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
Db       1640 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 1699

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
Db       1700 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 1753

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
Db       1754 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 1813

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
Db       1814 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 1873

Qy       1490 GCTCCATCC 1498
Db       1874 CCTCCTGCC 1882
```

RESULT 15
US-09-647-344A-14
; Sequence 14, Application US/09647344A
; Patent No. 6586180
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT.US
; CURRENT APPLICATION NUMBER: US/09/647,344A
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/06742
; PRIOR FILING DATE: 1999-03-28
; NUMBER OF SEQ ID NOS: 50
; SEQ ID NO 14
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pShuttle
US-09-647-344A-14

Query Match 3.1%; Score 49.4; DB 3; Length 8705;
Best Local Similarity 45.4%; Pred. No. 0.016;
Matches 222; Conservative 0; Mismatches 261; Indels 6; Gaps 1;

```
Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
Db       7265 CCTCCTCGACCCCGGCCTCCACTGCCTCCTCGACCCCGGCCTCCACCTCCTGCTCCTGCC 7324

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
Db       7325 CCTCCTGCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCT 7384

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC 1189
Db       7385 CCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCT 7444

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
Db       7445 CCTGCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCT 7504

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
Db       7505 GCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCT 7564

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
Db       7565 CCTGCCCCTCCTGCTCCTGCCCCT------CCTGCTCCTGCCCCTCCTGCCCCTCCTGCC 7618

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
Db       7619 CCTCCTCCTGCTCCTGCCCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTGCT 7678

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
Db       7679 CCTGCCCCTCCTCCTGCTCCTGCCCCTCCTGCCCCTCCTGCCCCTCCTCCTGCTCCTGCC 7738

Qy       1490 GCTCCATCC 1498
Db       7739 CCTCCTGCC 7747
```

Search completed: March 22, 2006, 03:02:01
Job time : 314 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 21:58:54 ; Search time 6.7 Seconds
(without alignments)
4566.706 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152020_26117/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_25_@runat_21032006_152020_26117 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2569	82.8	502	2	A48679	differentiated ker
2	341	11.0	1032	2	T34433	hypothetical prote
3	341	11.0	2232	2	T34434	hypothetical prote
4	337	10.9	13288	2	T03099	mucin, submaxillar
5	323.5	10.4	1428	2	T08852	lustrin A - Califo
6	318	10.3	1585	2	T31611	hypothetical prote
7	308.5	9.9	967	2	S66852	hypothetical prote
8	292.5	9.4	528	2	I47141	gastric mucin (clo
9	285	9.2	1188	2	S49915	extensin-like prot
10	284.5	9.2	1063	2	D86731	hypothetical prote
11	273	8.8	481	2	A35628	loricrin - mouse
12	271.5	8.8	1952	2	T48814	hypothetical prote
13	271	8.7	865	2	A47282	calcium-binding pr
14	271	8.7	1275	2	T33369	hypothetical prote
15	265.5	8.6	873	2	A47283	calphotin - fruit
16	262	8.4	286	2	C61615	sericin MG-2 - gre
17	261.5	8.4	1367	1	S48478	glucan 1,4-alpha-g
18	260.5	8.4	534	2	T39903	serine-rich protei
19	259	8.3	1459	2	T32271	hypothetical prote
20	258	8.3	1560	2	T02885	peroxisome prolife
21	249.5	8.0	1217	2	S52714	sericin1B - silkwo
22	248	8.0	2187	2	T30826	nascent polypeptid
23	246.5	7.9	3507	2	T34513	hypothetical prote
24	246	7.9	383	2	T46707	proteophosphoglyca
25	245	7.9	1932	1	A28490	DNA-directed RNA p
26	244	7.9	839	2	F75518	hypothetical prote
27	243	7.8	627	2	A44112	spidroin 2, dragli
28	241	7.8	2468	2	A83412	hypothetical prote
29	240	7.7	316	1	A38743	loricrin - human
30	239.5	7.7	786	2	T16509	hypothetical prote
31	238.5	7.7	536	2	T37544	hypothetical serin
32	237.5	7.7	683	2	A82704	1,4-beta-cellobios
33	237.5	7.7	805	2	T25795	hypothetical prote
34	237.5	7.7	1464	2	S59856	collagen alpha 1(I
35	236.5	7.6	467	2	A27677	DNA-directed RNA p
36	235.5	7.6	995	2	S50358	hypothetical prote
37	235.5	7.6	1896	1	RNFF2L	DNA-directed RNA p
38	235	7.6	1151	2	T18535	high molecular mas
39	234.5	7.6	406	2	S38170	SRP40 protein - ye
40	234.5	7.6	1970	1	S21054	DNA-directed RNA p
41	234.5	7.6	1970	2	I38186	RNA polymerase II
42	234	7.5	539	2	T28770	hypothetical prote
43	232.5	7.5	1589	2	T42233	submaxillary mucin
44	231	7.4	725	2	A41258	a-agglutinin core
45	230	7.4	1777	2	T34369	hypothetical prote

ALIGNMENTS

RESULT 1
A48679
differentiated keratinocyte S protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A48679
R;Zhou, Y.; Chaplin, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 9470-9474, 1993
A;Title: Identification in the HLA class I region of a gene expressed late in keratinoc
A;Reference number: A48679; MUID:94022396; PMID:8415725
A;Accession: A48679
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <ZHO>
A;Cross-references: UNIPROT:O43509; UNIPARC:UPI0000177376; GB:L20815
A;Note: sequence shown follows the authors' translation at positions 47 and 410
C;Genetics:
A;Map position: 6p21.3
C;Superfamily: loricrin

Alignment Scores:
Pred. No.: 3.8e-138 Length: 502
Score: 2569.00 Matches: 494
Percent Similarity: 99.4% Conservative: 3
Best Local Similarity: 98.8% Mismatches: 3
Query Match: 82.8% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x A48679 (1-502)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetLeuAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40
```

```
Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db        41 ProCysLysAspProThrCysIleThrSerProAsnAspProCysLeuThrGlyLysGly  60

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db        61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer  80

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db        81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db       101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db       121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db       141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db       161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db       181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db       201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db       221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db       241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db       261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db       281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db       301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db       321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db       341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db       361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db       381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db       401 SerSerSerIleSerSerSerSerGlyLeuProTyrHisProCysGlySerAlaSerGln 420

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500
```

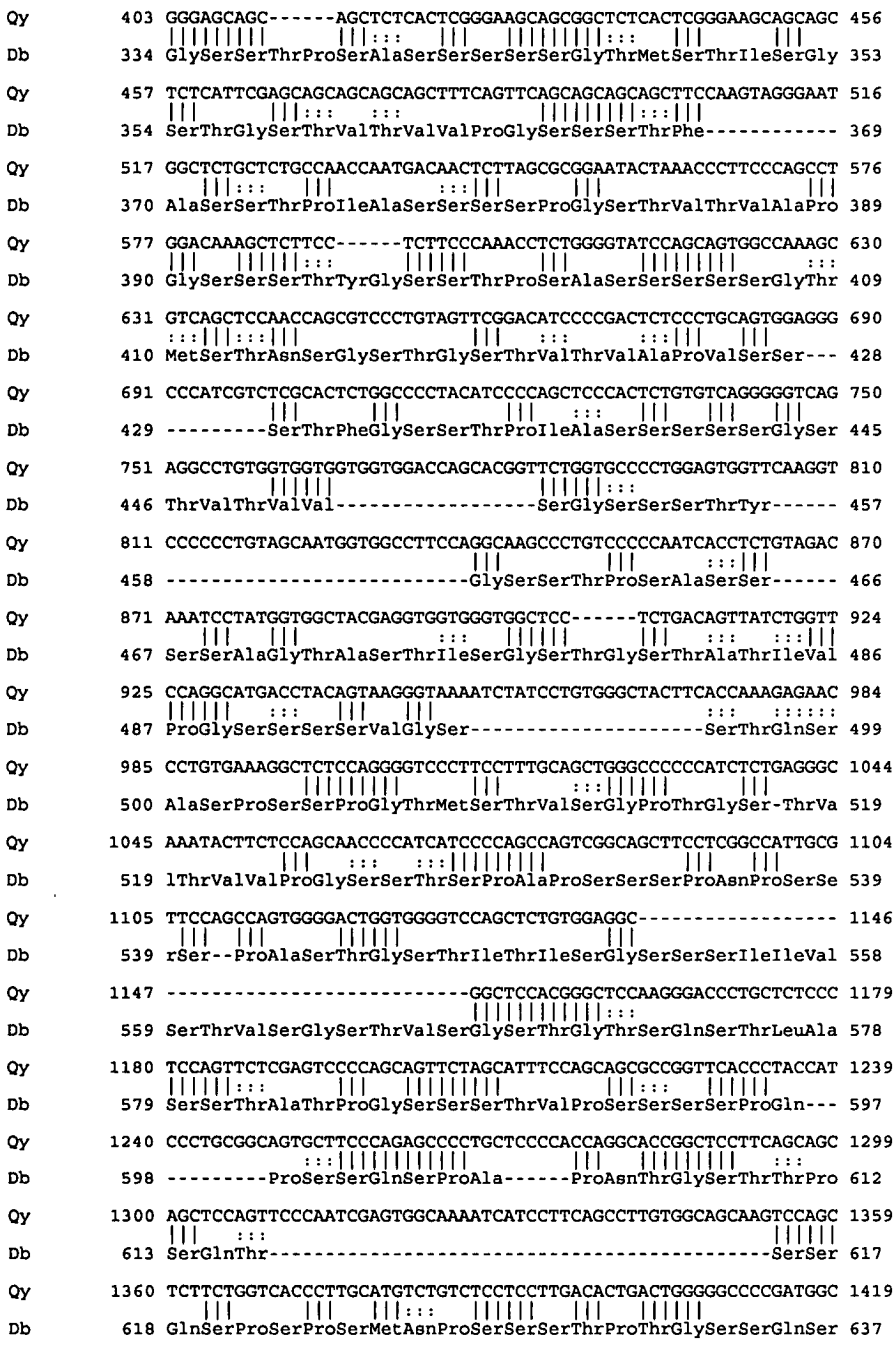
RESULT 2
T34433
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34433
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1032 <GEI>
A;Cross-references: UNIPROT:P91365; UNIPARC:UPI000007C995; EMBL:U80846; PIDN:AAC70889.1
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 839/1; 849/1; 875/3; 940/2; 966/2; 9

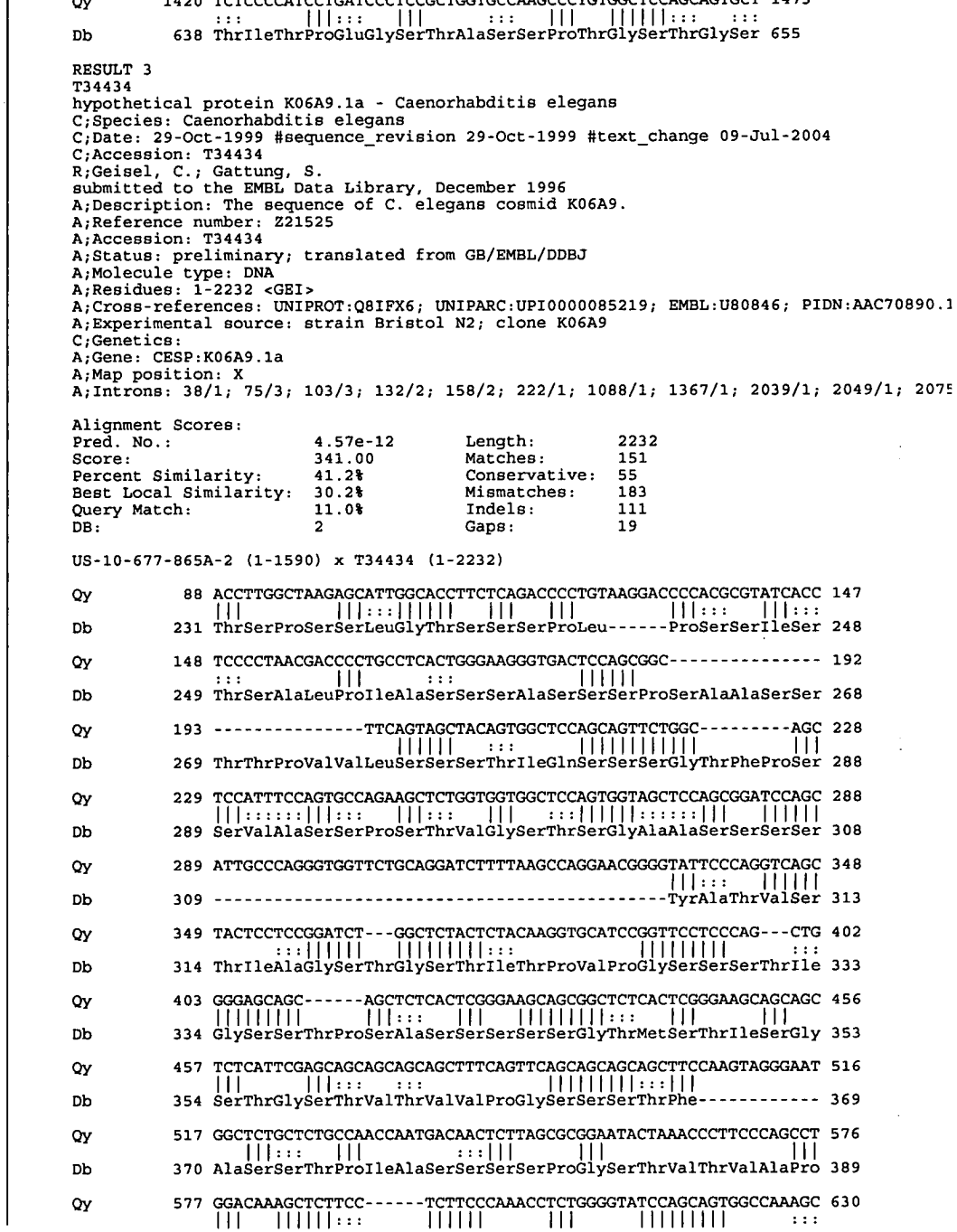
Alignment Scores:

Pred. No.:	4.8e-12	Length:	1032
Score:	341.00	Matches:	151
Percent Similarity:	41.2%	Conservative:	55
Best Local Similarity:	30.2%	Mismatches:	183
Query Match:	11.0%	Indels:	111
DB:	2	Gaps:	19

US-10-677-865A-2 (1-1590) x T34433 (1-1032)

```
Qy        88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
Db       231 ThrSerProSerSerLeuGlyThrSerSerSerProLeu------ProSerSerIleSer 248

Qy       148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGC--------------- 192
Db       249 ThrSerAlaLeuProIleAlaSerSerSerAlaSerSerSerProSerAlaAlaSerSer 268

Qy       193 ---------------TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGC---------AGC 228
Db       269 ThrThrProValValLeuSerSerSerThrIleGlnSerSerSerGlyThrPheProSer 288

Qy       229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
Db       289 SerValAlaSerSerProSerThrValGlySerThrSerGlyAlaAlaSerSerSerSer 308

Qy       289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
Db       309 ----------------------------------------------TyrAlaThrValSer 313

Qy       349 TACTCCTCCGGATCT---GGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG---CTG 402
Db       314 ThrIleAlaGlySerThrGlySerThrIleThrProValProGlySerSerSerThrIle 333
```





RESULT 3
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34434
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2232 <GEI>
A;Cross-references: UNIPROT:Q8IFX6; UNIPARC:UPI0000085219; EMBL:U80846; PIDN:AAC70890.1
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075

Alignment Scores:

Pred. No.:	4.57e-12	Length:	2232
Score:	341.00	Matches:	151
Percent Similarity:	41.2%	Conservative:	55
Best Local Similarity:	30.2%	Mismatches:	183
Query Match:	11.0%	Indels:	111
DB:	2	Gaps:	19

US-10-677-865A-2 (1-1590) x T34434 (1-2232)

```
Db        390 GlySerSerSerThrTyrGlySerSerThrProSerAlaSerSerSerSerSerGlyThr 409

Qy        631 GTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGG 690
              :::|||:::|||               |||   :::      :::|||   |||
Db        410 MetSerThrAsnSerGlySerThrGlySerThrValThrValAlaProValSerSer--- 428

Qy        691 CCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAG 750
                  |||      |||         |||   :::   |||   |||   |||
Db        429 ---------SerThrPheGlySerSerThrProIleAlaSerSerSerSerSerGlySer 445

Qy        751 AGGCCTGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGT 810
                  ||||||                ||||||:::
Db        446 ThrValThrValVal-----------------SerGlySerSerSerThrTyr------ 457

Qy        811 CCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGAC 870
                                        |||         |||      :::|||
Db        458 --------------------------GlySerSerThrProSerAlaSerSer------ 466

Qy        871 AAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCC------TCTGACAGTTATCTGGTT 924
              |||   |||            :::   ||||||      |||   :::   :::|||
Db        467 SerSerAlaGlyThrAlaSerThrIleSerGlySerThrGlySerThrAlaThrIleVal 486

Qy        925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984
              ||||||  :::   |||  |||                        :::   ::::::
Db        487 ProGlySerSerSerSerValGlySer-------------------SerThrGlnSer 499

Qy        985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
                      |||||||||      |||      :::||||||      |||
Db        500 AlaSerProSerSerProGlyThrMetSerThrValSerGlyProThrGlySer-ThrVa 519

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
                   |||   :::   :::|||||||||            |||   |||
Db        519 lThrValValProGlySerSerThrSerProAlaProSerSerSerProAsnProSerSe 539

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC----------------- 1146
              |||  |||      ||||||              |||
Db        539 rSer--ProAlaSerThrGlySerThrIleThrIleSerGlySerSerSerIleIleVal 558

Qy       1147 --------------------------GGCTCCACGGGCTCCAAGGGACCCTGCTCTCCC 1179
                                       ||||||||||||:::
Db        559 SerThrValSerGlySerThrValSerGlySerThrGlyThrSerGlnSerThrLeuAla 578

Qy       1180 TCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCAT 1239
              ||||||:::      |||   |||||||||         |||:::   ||||||
Db        579 SerSerThrAlaThrProGlySerSerSerThrValProSerSerSerSerProGln--- 597

Qy       1240 CCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGC 1299
                    :::||||||||||||         |||   |||||||||   :::
Db        598 ---------ProSerSerGlnSerProAla------ProAsnThrGlySerThrThrPro 612

Qy       1300 AGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGC 1359
              |||   :::                                              ||||||
Db        613 SerGlnThr-------------------------------------------SerSer 617

Qy       1360 TCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGC 1419
                 |||      |||   |||:::   ||||||   |||   ||||||
Db        618 GlnSerProSerProSerMetAsnProSerSerSerThrProThrGlySerSerGlnSer 637

Qy       1420 TCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
              :::      |||:::   |||     :::   |||   ||||||:::   :::
Db        638 ThrIleThrProGluGlySerThrAlaSerSerProThrGlySerThrGlySer 655

RESULT 4
T03099
mucin, submaxillary - pig
N;Alternate names: apomucin
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03099; A40009; A28528; B29789
R;Eckhardt, A.E.; Timpte, C.S.; DeLuca, A.W.; Hill, R.L.
J. Biol. Chem. 272, 33204-33210, 1997
A;Title: The complete cDNA sequence and structural polymorphism of the polypeptide chain
A;Reference number: Z14839; MUID:98070526; PMID:9407109
A;Accession: T03099
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-13288 <ECK>
A;Cross-references: UNIPROT:O18758; UNIPARC:UPI00001101C0; EMBL:AF005273; NID:g2581863;
R;Eckhardt, A.E.; Timpte, C.S.; Abernethy, J.L.; Zhao, Y.; Hill, R.L.
J. Biol. Chem. 266, 9678-9686, 1991
A;Title: Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain i
A;Reference number: A40009; MUID:91236743; PMID:2033060
A;Accession: A40009
A;Molecule type: mRNA
A;Residues: 12139-12167,'T',12169-13288 <EC3>
A;Cross-references: UNIPARC:UPI0000125C9F; GB:M61883; NID:g454837; PIDN:AAA30998.1; PID
R;Timpte, C.S.; Eckhardt, A.E.; Abernethy, J.L.; Hill, R.L.
J. Biol. Chem. 263, 1081-1088, 1988
A;Title: Porcine submaxillary gland apomucin contains tandemly repeated, identical sequ
A;Reference number: A28528; MUID:88087170; PMID:2826455
A;Accession: A28528
A;Molecule type: mRNA
A;Residues: 12139-12167,'T',12169-12641 <TIM>
A;Cross-references: UNIPARC:UPI000017920D; GB:M21174; GB:J03512; NID:g164321; PIDN:AAA3
A;Experimental source: submaxillary gland
R;Eckhardt, A.E.; Timpte, C.S.; Abernethy, J.L.; Toumadje, A.; Johnson Jr., W.C.; Hill,
J. Biol. Chem. 262, 11339-11344, 1987
A;Title: Structural properties of porcine submaxillary gland apomucin.
A;Reference number: A92606; MUID:87280230; PMID:3611111
A;Accession: B29789
A;Molecule type: protein
A;Residues: 1572-1607 <EC2>
A;Cross-references: UNIPARC:UPI000017920E
C;Superfamily: pig submaxillary mucin
C;Keywords: tandem repeat

Alignment Scores:
Pred. No.:              6.9e-12         Length:         13288
Score:                  337.00          Matches:        178
Percent Similarity:     39.2%           Conservative:   68
Best Local Similarity:  28.3%           Mismatches:     206
Query Match:            10.9%           Indels:         176
DB:                     2               Gaps:           31

US-10-677-865A-2 (1-1590) x T03099 (1-13288)

Qy          4 GGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCTG 63
              |||||||||   ||||||   :::         |||
Db       1576 GlySerSerGlyAlaProAlaValSerSer---GlyAlaSerGlnAlaAlaGlyThrSer 1594

Qy         64 CTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCC 123
                ||||||         |||||||||   |||   |||:::|||
Db       1595 GlyAlaGly---------ProGlyThrThrAlaSerSerValGly--------------- 1606

Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
                                    :::|||            |||   :::   |||   |||
Db       1607 -----------------ValThrGluThrAlaArgProSerValAlaGlySerGlyThr 1620

Qy        184 TCC---------AGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT--------------- 219
              :::         ||||||   |||       :::|||||||||||
Db       1621 ThrGlyThrValSerGlyAlaSerGlySerThrGlySerSerSerGlySerProGlyAla 1640

Qy        220 TCTGGCAGCTCCATT---------TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGT 270
              :::|||:::||||||           ::::::|||   |||      ||||||||||||
Db       1641 ThrGlyAlaSerIleGlyGlnProGluThrSerArgIleSerValAlaGlySerSerGly 1660

Qy        271 AGC---------TCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAG 321
              :::         |||||||||:::|||   |||   |||   |||
Db       1661 AlaProAlaValSerSerGlyAlaSerGlnAlaAlaGlyThrSerGlyAlaGly------ 1678

Qy        322 CCAGGAACGGGGTATTCCCAGGTCAGCTACTCC-------------------------- 354
              |||||||||      |||   |||      :::
Db       1679 ProGlyThrThrAlaSerSerValGlyValThrGluThrAlaArgProSerValAlaGly 1698

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
              ||||||:::      |||:::   ||||||||||||||||:::      ||||||||||||
```

```
Db       1699 SerGlyThrThrGlyThrValSerGlyAlaSerGlySerThr------GlySerSerSer 1716

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA-----------AGCAGCAGCTCTCAT 462
                 |||::::::|||:::      |||           |||      |||
Db       1717 GlySerProGlyAlaThrGlyAlaSerIleGlyGlnProGluThrSerArgIleServal 1736

Qy        463 TCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA---GTAGGGAATGGC 519
               :::   ||||||   :::         ||||||   :::   |||      |||
Db       1737 AlaGlySerSerGlyAlaProAlaValSerSerGlyAlaSerGlnAlaAlaGlyThrSer 1756

Qy        520 TCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCT--- 576
               |||   |||           |||      |||:::      ::::::|||
Db       1757 GlyAlaGlyProGlyThrThrAlaSerSerValGlyValThrGluThrAlaArgProSer 1776

Qy        577 --------------------------------GGACAAAGCTCTTCCTCTTCCCAAACC 603
                                               |||   |||   |||:::      :::
Db       1777 ValAlaGlySerGlyThrThrGlyThrValSerGlyAlaSerGlySerThrGlySerSer 1796

Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
               ||||||      |||   |||   :::   :::|||      :::|||   :::|||
Db       1797 SerGly------SerProGlyAlaThrGlyAlaSerIleGlyGlnProGluThrSerArg 1814

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCAC------------------ 705
               |||      :::      ||||||   |||   ||||||
Db       1815 IleSerValAlaGlySerSerGlyAlaProAlaValSerSerGlyAlaSerGlnAlaAla 1834

Qy        706 ------------TCTGGCCCC---TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAG 750
                         :::||||||            ||||||      |||:::
Db       1835 GlyThrSerGlyAlaGlyProGlyThrThrAlaSerSerValGlyValThrGluThrAla 1854

Qy        751 AGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGT 810
               ||||||   |||                 |||||||||      |||   |||   |||
Db       1855 ArgProSerValAla--------------GlySerGlyThrThrGlyThrValSerGly 1869

Qy        811 CCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGAC 870
                  |||   |||         |||   |||                   |||:::
Db       1870 AlaSerGlySerThrGlySerSerSerGlySerProGlyAlaThrGlyAlaSerIleGly 1889

Qy        871 AAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT------ 924
               :::                  |||   |||||||||   :::      |||
Db       1890 GlnProGluThrSerArgIleSerValAlaGlySerSerGlyAlaProAlaValSerSer 1909

Qy        925 -----------------------------------CCAGGCATGACCTACAGTAAGGGT 948
                                                  ||||||   |||   |||
Db       1910 GlyAlaSerGlnAlaAlaGlyThrSerGlyAlaGlyProGlyThrThrAlaSerSer--- 1928

Qy        949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCT--------------------- 987
                     ||||||            |||
Db       1929 ------------ValGlyValThrGluThrAlaArgProSerValAlaGlySerGlyThr 1944

Qy        988 ---------GTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCC------ 1032
                  |||   |||:::   |||           ::::::|||   |||
Db       1945 ThrGlyThrValSerGlyAlaSerGlySerThrGlySerSerSerGlySerProGlyAla 1964

Qy       1033 ------ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCA 1086
                     |||   |||:::   :::|||      :::          |||
Db       1965 ThrGlyAlaSerIleGlyGlnProGluThrSerArgIleSerValAlaGlySerSerGly 1984

Qy       1087 GCTTCC------TCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTC--------- 1131
               |||         |||:::   |||   |||      ||||||   |||
Db       1985 AlaProAlaValSerSerGlyAlaSerGlnAlaAlaGlyThrSerGlyAlaGlyProGly 2004

Qy       1132 ----------------------------------------CAGCTCTGTGGAGGCGGC 1149
                                                        :::   |||   |||
Db       2005 ThrThrAlaSerSerValGlyValThrGluThrAlaArgProSerValAlaGlySerGly 2024

Qy       1150 TCCACG--------------------GGCTCCAAGGGACCCTGCTCT------------ 1176
               :::|||                  ||||||   |||      |||
Db       2025 ThrThrGlyThrValSerGlyAlaSerGlySerThrGlySerSerSerGlySerProGly 2044

Qy       1177 ----------------------CCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATT 1212
                                     |||   :::||||||:::            |||:::
Db       2045 AlaThrGlyAlaSerIleGlyGlnProGluThrSerArgIle------------SerVal 2060

Qy       1213 TCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCC 1272
               :::   |||:::|||:::|||           |||::::::      :::      |||
Db       2061 AlaGlySerSerGlyAlaProAlaValSerSerGlyAlaSerGlnAlaAlaGlyThrSer 2080

Qy       1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGC---------------------------- 1302
                  |||   |||:::   :::||||||
Db       2081 GlyAlaGlyProGlyThrThrAlaSerSerValGlyValThrGluThrAlaArgProSer 2100

Qy       1303 ---TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGC 1359
                  :::   |||   ::::::|||      :::             ||||||      |||
Db       2101 ValAlaGlySerGlyThrThrGly---ThrValSerGlyAlaSerGlySerThrGlySer 2119

Qy       1360 TCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGAT--- 1416
               |||||||||   |||      :::  :::|||:::         |||   |||:::
Db       2120 SerSerGlySerProGlyAlaThrGlyAlaSerIle---------GlyGlnProGluThr 2136

Qy       1417 --------------------GGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC---AAG 1452
                                   |||      |||      |||:::||||||   :::
Db       2137 SerArgIleSerValAlaGlySerSerGlyAlaProAlaValSerSerGlyAlaSerGln 2156

Qy       1453 CCCTGTGGCTCCAGCAGTGCTGGA 1476
                  |||:::|||   ||||||
Db       2157 AlaAlaGlyThrSerGlyAlaGly 2164
```

RESULT 5
T08852
lustrin A - California red abalone
C;Species: Haliotis rufescens (California red abalone)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08852
R;Shen, X.; Belcher, A.M.; Hansma, P.K.; Stucky, G.D.; Morse, D.E.
J. Biol. Chem. 272, 32472-32481, 1997
A;Title: Molecular cloning and characterization of lustrin A, a matrix protein from sh
A;Reference number: Z16496; MUID:98070424; PMID:9405458
A;Accession: T08852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1428 <SHE>
A;Cross-references: UNIPROT:O44341; UNIPARC:UPI00000836D6; EMBL:AF023459; NID:g2723361,
A;Experimental source: tissue type mantle (shell and pearl nacre); cell type pallial
C;Keywords: extracellular matrix; extracellular protein
F;1382-1426/Domain: antileukoproteinase repeat homology <ALP>

Alignment Scores:

Pred. No.:	4.6e-11	Length:	1428
Score:	323.50	Matches:	158
Percent Similarity:	37.1%	Conservative:	37
Best Local Similarity:	30.0%	Mismatches:	162
Query Match:	10.4%	Indels:	169
DB:	2	Gaps:	25

US-10-677-865A-2 (1-1590) x T08852 (1-1428)

```
Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACG---------------------------- 138
               |||      ||||||      ||||||
Db        916 ThrCysValAspProSerProProProThrArgLysCysProValGlyLysProValLeu 935

Qy        139 ------CGTATCACCTCC-----------------------------CCTAACGACCCC 162
                   |||:::|||                              |||   |||
Db        936 ThrProArgLeuThrGluPheArgCysTyrProArgValArgLeuCysProGlyAspSer 955

Qy        163 ---TGCCTCACTGGGAAGGGTGACTCCAGCGGCTTC------------------------ 195
                  ||||||   |||   ||||||      |||
Db        956 PheCysLeuArgGlyProGlyAspGluProGlyValCysCysTrpAspAsnArgLeuArg 975

Qy        196 ---------AGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGA 246
                        |||   |||||||||   |||         |||   ||||||:::
Db        976 ProThrGlnGlySerGlySerGlySerGlySerGlySerGlySerGlySerSerSerSer 995

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCT 306
               ||||||   |||   :::|||||||||   |||||||||           |||   |||
Db        996 SerSerSerGlySerThrSerGlySerGlySerGlySer------GlySerGlySerSer 1013
```

```
Qy       307 GCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGC 366
             :::||||||         |||:::|||   |||            :::|||||||||||||
Db      1014 SerGlySer---------GlySerGlySerSerSer---------AlaSerGlySerGly 1027

Qy       367 TCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGA 426
             |||:::         :::||||||||||||         ||||||   |||   ||||||
Db      1028 SerSerSerGlySerSerSerGlySerSerSerGlySerSerSerGlySerSerSerGly 1047

Qy       427 AGC---AGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 483
             |||   |||||||||   |||||||||||||||:::      |||   |||||||||
Db      1048 SerGlySerGlySerGlySerGlySerSerSerAlaSerGlySerGlySerSerSerGly 1067

Qy       484 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 543
                  |||   ||||||       |||:::||||||:::      :::
Db      1068 SerSerSerGlySerSerSerGlySerGlySerGlySerSerSerAlaSerGlySerGly 1087

Qy       544 TCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACC 603
             |||      |||             |||      |||   ||||||   |||
Db      1088 SerSerSerGly-----------SerGlySerGlySerSerSerGlySerGlySerGly 1103

Qy       604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
             ||||||     ||||||      |||   |||   :::
Db      1104 SerGlySerGlySerSerSerGlySerGlySerGlySerSer----------------- 1117

Qy       664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
                             ||||||      ||||||:::|||
Db      1118 -----------------SerGlySerSerSerValSerAsnSerTrp---------Thr 1128

Qy       724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
                |||   |||   ||||||                                     |||
Db      1129 GlySerGlySerSerSerGlySer---------------------------------Gly 1137

Qy       784 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
             ||||||:::              |||         |||:::|||
Db      1138 SerGlySerSerSer---TrpSerGlySerGlySerSerSerGly--------------- 1151

Qy       844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
                                                                     |||
Db      1152 -----------------------------------------------------ThrGlySer 1154

Qy       904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
                |||   |||:::         |||   :::         |||
Db      1155 GlySerSerSerTrpPheGlySerGlySerSerSerGlySerGly--------------- 1169

Qy       964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
                                                 |||           |||   |||:::
Db      1170 --------------------------SerAspSerSerSerGlySerSerSerAlaSer 1180

Qy      1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
             |||         |||   |||                                 |||   |||
Db      1181 GlySerGlySerSerSerGlySer-----------------------SerSerGlySer 1192

Qy      1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143
                :::   |||::::::   |||         |||:::|||                 |||
Db      1193 GlySerGlySerSerLeuTrpPhe---------GlySerGlySerSerSerGlyThrGly 1209

Qy      1144 GGCGGCTCCACG------GGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCC 1197
                ||||||:::          |||   |||       |||   |||||||||
Db      1210 SerGlySerSerSerGlySerSerSerGlySerGlySerAspSerSerSerGlySerSer 1229

Qy      1198 AGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCC 1257
             |||   |||:::      ||||||:::||||||                 ||||||||||||
Db      1230 SerGlySerThrSerGlySerSerSerGlySer--------------GlySerAlaSer 1244

Qy      1258 CAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCG 1317
                |||                |||||||||||         :::|||
Db      1245 GlySer--------------GlyThrGlySerGlyLysGlyAlaSerTyrAspThrAsp 1259

Qy      1318 AGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCAC------ 1371
             :::                      ||||||   :::   |||   |||:::
Db      1260 AlaAspSer-----------------GlySerAspAsnArgSerProGlyTyrLeuPro 1273

Qy      1372 ------CCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCC------------ 1413
                   ||||||                  :::      |||   |||
Db      1274 GlnAspProCysThrProGlyLeuTyrIleAsnCysThrAlaGlyThrCysArgLeuThr 1293

Qy      1414 ----------------------------------------------GATGGCTCTCCCCAT 1428
                                                           |||   ||||||
Db      1294 AlaTrpCysLeuTyrAsnPheCysProAlaValProThrCysValAspSerSerPro--- 1312

Qy      1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGA------------ 1476
                     ::::::|||      |||   |||      :::
Db      1313 ------AspAlaSerGlyGluCysProValGlyLeuProAlaLeuAsnTyrPheAsnLys 1330

Qy      1477 AAGATCCCCTGCCGCTCC 1494
             ::::::   ||||||:::
Db      1331 GluValSerCysArgThr 1336
```

RESULT 6
T31611
hypothetical protein Y50E8A.g - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31611
R;Steward, C.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21047
A;Accession: T31611
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1585 <WIL>
A;Cross-references: UNIPARC:UPI000017BC9F; EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1
A;Experimental source: clone Y50E8A
C;Genetics:
A;Gene: CESP:Y50E8A.g
A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Alignment Scores:

Pred. No.:	9.36e-11	Length:	1585
Score:	318.00	Matches:	173
Percent Similarity:	31.9%	Conservative:	29
Best Local Similarity:	27.3%	Mismatches:	217
Query Match:	10.3%	Indels:	215
DB:	2	Gaps:	22

US-10-677-865A-2 (1-1590) x T31611 (1-1585)

```
Qy       107 GCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCC 166
             ||||||          |||          |||          ||||||          ||||||
Db       970 AlaProProProProProProProProProProProProProProProProAlaProAla 989

Qy       167 TCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCA 226
                                 ||||||   :::   :::  ||||||       |||   ||||
Db       990 -----------------ProAlaProAlaProSer--SerGlyGlyTyrSerGlyGlyS 1003

Qy       227 GCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCA 286
             |||||      ||||||        ||||||||||||||||||     ||||||    |
Db      1003 erSerGlyGlySerAlaAlaGlyGlyGlyGlySerSerGlyGlyTyrSerGlyGlyS 1023

Qy       287 GCATTGCCCAG---------------------------------------------- 297
             ||   |||
Db      1023 erAlaAlaProProProProProProProProProProAlaProAlaProAlaProAlaP 1043

Qy       298 ----------------------GGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGT 334
                                    |||   |||   ||||||       |||   |||
Db      1043 roSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyG 1063

Qy       335 ATTCCCAGGTCAGCTACTCCTCCGGA-------------------------------- 360
               |||         ||||||   |||
Db      1063 lySerSerGlyGlyTyrSerGlyGlySerThrAlaProProProProProProProP 1083

Qy       361 ----------------------------------------TCTGGCTCTACTCTACAAG 379
                                                    ||||||           |
Db      1083 roProProProProAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSerGlyG 1103
```

```
Qy       380 GTGCATCCGGTTCCTCCCAGCTG-----------GGGAGCAGCAGCTCTCACTCGGGA- 426
Db      1103 lySerSerGlyGlySerAlaAlaGlyGlyGlyGlySerSerGlyGlyTyrSerGlyG 1123

Qy       426 ------------------------------------------------------------ 426

Db      1123 lySerAlaAlaProProProProProProProProProProAlaProAlaProAlaProA 1143

Qy       427 -----AGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCT 481
Db      1143 laProSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGly- 1162

Qy       482 TTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG------------ 528
Db      1163 --------GlyGlySerSerGlyGlyTyrSerGlyGlySerAlaAlaProProProProP 1180

Qy       529 --------CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGAC 580
Db      1180 roProProProProProProProProAlaProAlaProAlaProAlaProSerSerGlyGlyT 1200

Qy       581 AAAGCTCTTCCTCTTCCCAAACCTCT--------------GGGGTATCCAGCAGTGGCC 625
Db      1200 yrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyGlyGlySerSerGlyGlyT 1220

Qy       626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
Db      1220 yrSerGlyGlySerAlaAlaProProProProProProProProProProProAlaP 1240

Qy       686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGG 745
Db      1240 roAlaProAlaProAla-----------------ProSerSerGlyGlyTyrSerGlyG 1254

Qy       746 GTCA----------GAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT------------ 783
Db      1254 lySerSerGlyGlySerAlaAlaGlyGlyGlyGlySerSer-GlyGlyTyrThrGly 1273

Qy       784 ---TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db      1274 GlySerAlaAlaProProProProProProPro-----------------ProPro 1287

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db      1288 ProProProAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSer------Gly 1305

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db      1306 GlySerSerGlyGlySerAlaAlaGlyGlyGlyGlyGlySerSerGlyGly---TyrSer 1324

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db      1325 GlyGlySerAlaAlaProProProProProProProProProProProAlaProAla 1344

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAAC----------------- 1062
Db      1345 ProAlaProAlaProSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAla 1364

Qy      1062 ------------------------------------------------------------ 1062

Db      1365 GlyGlyGlyGlyGlySerSerGlyGlyTyrThrGlyGlySerAlaAlaProProProPro 1384

Qy      1063 --------------------CCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATT 1101
Db      1385 ProProProProProProProProAlaProAlaProAlaProAlaProSerSerGly 1404

Qy      1102 GCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCC 1161
Db      1405 GlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGlyGlyGlyGly------GlySer 1422

Qy      1162 AAGGGA-----------------------------------CCCTGCTCTCCCTCCAGT 1185
Db      1423 SerGlyGlyTyrSerGlyGlySerAlaAlaProProProProProProAlaProAlaPro 1442

Qy      1186 TCTCGAGTCCCCAGCAGTTCTAGCATTTCC------AGCAGCGCCGGTTCA--------- 1230

Db      1443 AlaProAlaProSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGly 1462

Qy      1230 ------------------------------------------------------------ 1230

Db      1463 GlyGlyGlyGlySerSerGlyGlyTyrThrGlyGlySerAlaAlaProProProProPro 1482

Qy      1231 CCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCC 1290
Db      1483 ProProProProProProProProAlaProAlaProAlaProAlaProSerSerGlyGly 1502

Qy      1291 TTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
Db      1503 TyrSerGlyGlySerSerGlyGlySerAlaAlaGly-----------------GlyGly 1516

Qy      1351 AAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC 1410
Db      1517 GlyGlySerSerGlyGlyTyr----------------------------ThrGlyGly 1526

Qy      1411 CCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGT 1470
Db      1527 SerAlaAlaProProProProPro-ProProProProProProProProAlaProAla-- 1545

Qy      1471 GCTGGAAAGATCCCCTGCCGCTCCATCCGGG 1501
Db      1546 -------ProAlaProAlaProSerSerGly 1553
```

RESULT 7
S66852
hypothetical protein YOL155c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O0419; protein AOF1001
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66852; S66854; S67325; S70380
R;Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas,
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66814
A;Accession: S66852
A;Molecule type: DNA
A;Residues: 1-967 <ARI>
A;Cross-references: UNIPROT:Q08294; UNIPARC:UPI000006C5C2; EMBL:Z74897; NID:g1420063; P
A;Experimental source: strain S288C
R;Gaillon, L.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66854
A;Accession: S66854
A;Molecule type: DNA
A;Residues: 1-967 <GAI>
A;Cross-references: UNIPARC:UPI000006C5C2; EMBL:Z74897; NID:g1420063; PID:e251930; PID:
A;Experimental source: strain S288C
R;Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Aldea, M.; Casas, C.; Ario, J.; Herrero, E
submitted to the EMBL Data Library, July 1995
A;Description: Analysis of the DNA sequence of a 15500 bp fragment of the left arm of ch
pen reading frames.
A;Reference number: S67324
A;Accession: S67325
A;Molecule type: DNA
A;Residues: 1-164,'STSITSGSSSATESGSSVSGSTSATESGSSASGSS',166-186,'V',188-967 <GAM>
A;Cross-references: UNIPARC:UPI0000069FE4; EMBL:X89715; NID:g1177620; PID:e190152; PID:
R;Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Arino, J.; Aldea, M.; Casas, C.; Herrero, E
Yeast 12, 709-714, 1996
A;Title: Analysis of the DNA sequence of a 15,500 bp fragment near the left telomere of
and two new open reading frames.
A;Reference number: S70379; MUID:96405919; PMID:8810044
A;Accession: S70380
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 55-164,'STSITSGSSSATESGSSVSGSTSATESGSSASGSS',166-186,'V',188-286;745-773;781
A;Cross-references: UNIPARC:UPI000017CCC3; UNIPARC:UPI000017CCC4; UNIPARC:UPI000017CCC5;
C;Genetics:
A;Cross-references: SGD:S0005515
A;Map position: 15L
A;Note: YOL155c

Alignment Scores:

Pred. No.:	3.33e-10	Length:	967
Score:	308.50	Matches:	158
Percent Similarity:	37.9%	Conservative:	65
Best Local Similarity:	26.8%	Mismatches:	209
Query Match:	9.9%	Indels:	157
DB:	2	Gaps:	24

US-10-677-865A-2 (1-1590) x S66852 (1-967)

```
Qy        133 CCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCC---AGC 189
Db         83 ProSerThrSerSerSerGluValSerSerSerIleThrSerSerGlySerSerValSer 102

Qy        190 GGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGC 249
Db        103 GlySerSerSerIleThrSerSerGlySerSerValSerSerSerSerSerAlaThrGlu 122

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCC------------AGCGGATCCAGCATTGCCCAG 297
Db        123 SerGlySerSerAlaSerGlySerSerSerAlaThrGluSerGlySerSerValSer--- 141

Qy        298 GGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCC 357
Db        142 ---GlySerSerThrSerIleThrSerGlySerSerSerAlaThrGluSerGlySerSer 160

Qy        358 GGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC--- 414
Db        161 ValSerGlySerThrSerAlaThrGluSerGlySerSerAlaSerGlySerSerSerAla 180

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGC---------------------- 450
Db        181 ThrGluSerGlySerSerAlaSerGlySerSerSerAlaThrGluSerGlySerServal 200

Qy        451 ------AGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
Db        201 SerGlySerSerSerAlaThrGluSerGlySerSerValSerGlySerSerSerAlaThr 220

Qy        505 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAAC 564
Db        221 GluSerGlySerAlaSerSerValProSerSerSerGlySer------------------ 234

Qy        565 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCT------------------ 606
Db        235 ValThrGluSerGlySerSerSerSerAlaSerGluSerSerIleThrGlnSerGlyThr 254

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
Db        255 AlaSerGlySerSerAlaSerSerThrSerGlySerValThrGlnSerGlySerSerVal 274

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
Db        275 SerGlySerSerAlaSerSerAlaProGlyIleSerSerSerIleProGlnSerThrSer 294

Qy        727 TCCCACTCTGTGTCAGGGGGTCAG------AGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db        295 SerAlaSerThrAlaSerGlySerIleThrSerGlyThrLeuSerSerIleThrSerSer 314

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db        315 AlaSerSerAlaThrAlaThrAlaSerAsnSerLeuSerSerSerAspGlyThrIleTyr 334

Qy        841 AAGCCCTGTCCCCCAATCACC---------------------------------------- 861
Db        335 LeuProSerThrThrIleSerGlyAspIleThrLeuThrGlySerValIleAlaThrGlu 354

Qy        862 TCTGTAGACAAATCCTATGGTGGC---TACGAGGTGGTGGGTGGC--------------- 903
Db        355 AlaValGluValAlaAlaGlyGlyLysLeuThrLeuLeuAspGlyAspLysTyrValPhe 374

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAA---ATCTATCCT 960
Db        375 SerAlaAspPheIleIleHisGlyGlyValPheValGluLysSerLysProThrTyrPro 394

Qy        961 ----------------------------------------GTGGGCTACTTCACCAAAGAG 981
Db        395 GlyThrGluPheAspIleSerGlyGluAsnPheAspValSerGlyThrPheAsnAlaGlu 414

Qy        982 AACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGG--------------- 1026
Db        415 GluProAlaAlaSerSerAlaSerAlaTyrSerPheThrProGlySerPheAspAsnSer 434

Qy       1027 ------------------------------------------------------CCCCCC 1032
Db        435 GlyAspIleSerLeuSerLeuSerGluSerThrLysGlyGluValThrPheSerProTyr 454

Qy       1033 ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT--- 1089
Db        455 SerAsnSerGlyAlaPheSerPheSerAsnAlaIleLeuAsnGlyGlySerValSerGly 474

Qy       1089 ------------------------------------------------------------ 1089
Db        475 LeuGlnArgArgAlaGluSerGlySerValAsnAsnGlyGluIleAsnLeuGluAsnGly 494

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTG---GGGACTGGTGGGGTCCAGCTCTGTGGAGGC 1146
Db        495 SerThrTyrValValValGluProValSerGlySerGlyThrIleAsnIleIleSerGly 514

Qy       1147 GGC------------------------TCCACGGGCTCC------------AAGGGA--- 1167
Db        515 AsnLeuTyrLeuHisTyrProAspThrPheThrGlyGlnThrValValPheLysGlyGlu 534

Qy       1168 -----------------CCCTGCTCTCCCTCCAGTTCTCGAGTCCCC------------ 1197
Db        535 GlyValLeuAlaValAspProThrGluThrAsnThrThrProIleProValValGlyTyr 554

Qy       1198 ----------------------------------AGCAGTTCTAGCATTTCCAGCAGC 1221
Db        555 ThrGlyGluAsnGlnIleAlaIleThrAlaAspValThrAlaLeuSerTyrAspSerAla 574

Qy       1222 GCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGC 1281
Db        575 ThrGlyValLeuThrAlaThrGlnGlyAsnSerGlnPheSer------PheSerIleGly 592

Qy       1282 ACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC------------ 1329
Db        593 ThrGly---PheSerSerSerGlyPheAsnValSerGluGlyThrPheAlaGlyAlaTyr 611

Qy       1330 ---ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTC 1386
Db        612 AlaTyrTyrLeuAsnTyrGlyGlyValValAlaSerSerAlaThrPro------------ 627

Qy       1387 TCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGT 1446
Db        628 SerSerThrSerThrThrSerGlyAlaThrAsnSerThrSerGlySerThrSerPheGly 647

Qy       1447 GCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
Db        648 AlaSerValThrGlySerThrAlaSer 656
```

RESULT 8
I47141
gastric mucin (clone PGM-2A) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: I47141; S55315
R;Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
Gastroenterology 106, 200, 1994
A;Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a novel
A;Reference number: I47141; MUID:94102478; PMID:7506218
A;Accession: I47141
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-528 <TUR>
A;Cross-references: UNIPROT:Q29071; UNIPARC:UPI0000086746; EMBL:U10281; NID:g915205; PI

```
R;Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
Biochem. J. 308, 89-96, 1995
A;Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
A;Reference number: S55315; MUID:95275264; PMID:7755593
A;Accession: S55315
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-528 <TU2>
A;Cross-references: UNIPARC:UPI0000086746; GB:U10281; NID:g915205; PIDN:AAC48526.1; PID:
C;Superfamily: pig submaxillary mucin

Alignment Scores:
Pred. No.:              2.78e-09      Length:         528
Score:                  292.50        Matches:        151
Percent Similarity:     38.8%         Conservative:   58
Best Local Similarity:  28.0%         Mismatches:     223
Query Match:            9.4%          Indels:         107
DB:                     2             Gaps:           24

US-10-677-865A-2 (1-1590) x I47141 (1-528)

Qy         82 CCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC---------CCCTGTAAGGAC 132
Db         45 ProThrThrSerAlaThrSerValGlnThrSerSerSerSerSerProProIleSerSer 64

Qy        133 -----------------------------------CCCACGCGTATCACCTCCCCTAAC 156
Db         65 ThrIleSerValGlnThrSerSerSerSerSerValProThrThrSerThrThrSerVal 84

Qy        157 GACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGC 216
Db         85 GlnProSerSerSerSerSerAlaProThrThrArgAlaThrSerValGlnSerSerSer 104

Qy        217 AGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGC--- 273
Db        105 SerSerSerAlaProIleSerSerThrThrSerValGlnProSerSerSerGlySerVal 124

Qy        274 ------TCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
Db        125 ProThrThrSerAlaThrSerValGln------SerSerSerSerSerSerAlaProThr 142

Qy        328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCC 387
Db        143 ThrSerAlaThrSerValGlnProSerSerSerSerSerProProIle------------ 158

Qy        388 GGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA 447
Db        159 ---SerSerThrValSerValGlnProSerSerSerSerSerAlaProThrThrSerAla 177

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507
Db        178 ThrSerValGlnProSerSerSerSerSerProProIleSerSerThrValSerValGln 197

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567
Db        198 ThrSerSerSerSerSerValProThrThrSerThrThr-----------------Ser 211

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627
Db        212 ValGlnProSerSerSerSerSerValProThrThrSerAlaThrSerValArgSerSer 231

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687
Db        232 SerSerSerSerThrProIleProSerThrThrSerValGlnProSerSerSerSerSer 251

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGT 747
Db        252 AlaProThrThrSerAlaThrSerValGln---ProSerSerSerSer------------ 266

Qy        748 CAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAA 807
Db        267 SerThrProIle--------------------------------ProSerThrThrSer 275

Qy        808 GGTCCCCCCTGTAGCAATGGTGGCCTTCCA---------------------------- 837
Db        276 ValGlnProSerSerSerSerSerAlaProThrThrSerAlaThrSerValGlnProSer 295

Qy        838 GGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTG 897
Db        296 SerSerSerSerProProIleSerSerThr-----------------IleSerValGln 309

Qy        898 GGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAT 957
Db        310 ProSerSerSerSerSer-----------SerProThrThrSerThrThrSerValGln 325

Qy        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC---CCTTCC 1014
Db        326 ProSerSer------SerGlySerAlaProThrThrSerAlaThrSerValGlnProSer 343

Qy       1015 TTTGCAGCTGGGCCCCCCATCTCT--------------GAGGGCAAATACTTCTCCAGC 1059
Db        344 SerSerSerSerProProIleSerSerThrIleSerValGlnProSerSerSerSerSer 363

Qy       1060 AACCCC-----------------ATCATCCCCAGCCAGTCGGCAGCT---------TCC 1092
Db        364 SerProThrThrSerThrThrSerValGlnProSerSerSerGlySerAlaProThrThr 383

Qy       1093 TCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCC 1152
Db        384 SerAlaThrSerValGlnProSerSerSerSerSerValProThrThrSerAlaThrSer 403

Qy       1153 ACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATT 1212
Db        404 ValArgSerSer---------SerSerSerSerThrProIleProThrThrThrSerVal 420

Qy       1213 TCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT---------------TCC 1257
Db        421 GlnProSerSerSerSerSerVal---ProThrThrSerAlaThrSerValGlnThrSer 439

Qy       1258 CAGAGCCCCTGCTCCCCA---CCAGGCACCGGCTCCTTC------AGCAGCAGCTCCAGT 1308
Db        440 SerSerSerSerThrProIleProSerThrThrSerValGlnProSerSerSerSerSer 459

Qy       1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
Db        460 AlaProThrThrSerAlaThrSerValGlnPro-----------SerSerSerSerSer 475

Qy       1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
Db        476 ProProIleSerSerThrIleSerVal--------------GlnProSerSerSerSer 490

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
Db        491 SerSerProThrThrSerThrThrSerValGlnProSerSerSerGlySerAlaProThr 510

Qy       1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGC 1541
Db        511 ThrSerAlaThrSerVal--GlnProSerSerSerSerProProIleSer 527

RESULT 9
S49915
extensin-like protein - maize
C;Species: Zea mays (maize)
C;Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: S49915
R;Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.
submitted to the EMBL Data Library, June 1994
A;Description: Pex genes: pollen-specific genes with extensin-like domains.
A;Reference number: S49915
A;Accession: S49915
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1188 <RUB>
A;Cross-references: UNIPROT:Q41805; UNIPARC:UPI00000A1616; EMBL:Z34465; NID:g600117; PI

Alignment Scores:
```

Pred. No.:	7.04e-09	Length:	1188
Score:	285.00	Matches:	135
Percent Similarity:	36.3%	Conservative:	60
Best Local Similarity:	25.1%	Mismatches:	191
Query Match:	9.2%	Indels:	151
DB:	2	Gaps:	23

US-10-677-865A-2 (1-1590) x S49915 (1-1188)

```
Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCA 169
Db        724 ProValSerSerProProGlnThrProLysSerSerProPro---ProAlaProValSer 742

Qy        170 CTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCT 229
Db        743 SerProProProThrProValSerSerProProAlaLeuAlaPro------ValSerSer 760

Qy        230 CCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCA 289
Db        761 ProProSerValLysSerSerProProProAlaProLeuSerSerPro------ProPro 778

Qy        290 TTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCT 349
Db        779 AlaProGlnVal-----------------------------------------LysSerSer 785

Qy        350 ACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCA 409
Db        786 ProProPro--------------ValGlnValSerSerProProPro------------ 796

Qy        410 GCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCA 469
Db        796 ------------------------------------------------------------ 796

Qy        470 GCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGC 529
Db        797 ------AlaProLysSerSer---------------------------------------- 801

Qy        530 CAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTT 589
Db        802 ProProLeuAlaProValSerSerPro------------ProGlnValGluLysThrSer 817

Qy        590 CCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTC 649
Db        818 ProProProAlaProLeuSerSerProProLeuAlaProLysSerSerProProHisVal 837

Qy        650 ------------------CCTGTAGTTCGGACATCC------------------------ 667
Db        838 ValValSerSerProProProValValLysSerSerProProProAlaProValSerSer 857

Qy        668 ---CCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCT--------- 715
Db        858 ProProLeuThrProLysProAlaSerProProAlaHisValSerSerProProGluVal 877

Qy        716 ------------ACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGG 763
Db        878 ValLysProSerThrProProAlaPro------------------------------- 886

Qy        764 TGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCA 823
Db        887 ---------ThrThrValIleSerProProSerGluProLysSerSerProProProThr 903

Qy        824 ATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT---------CTGTAGACAAAT 874
Db        904 ProValSerLeuProProProIleValLysSerSerProProProAlaMetValSerSer 923

Qy        875 CCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGA 934
Db        924 ProProMetThrProLys------SerSerProProProValValValSerSerProPro 941

Qy        935 CCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAG 994
Db        942 ProThrValLysSerSerProProProAlaProValSerSerProProAlaThrProLys 961

Qy        995 GCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCT 1054
Db        962 SerSerProProProAlaProValAsnLeuProProProGluValLysSerSerProPro 981

Qy       1055 CCA------------GCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
Db        982 ProThrProValSerSerProProProAlaProLysSerSerProProProAlaProMet 1001

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCC--------------------AGCTCTGTG 1141
Db       1002 SerSerProProProProGluValLysSerProProProProAlaProValSerSerPro 1021

Qy       1142 GAGGCGGCTCCACGGGCTCCAAGGGACCCTGCT------------CTCCCTCCAGTTCTC 1189
Db       1022 ProProProValLysSerProProProProAlaProValSerSerProProProProVal 1041

Qy       1190 GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCA 1249
Db       1042 LysSerProPro------------ProProAlaProValSerSerProProProProVal 1057

Qy       1250 GTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTT 1309
Db       1058 LysSerProProProProAlaProIleSerSerProProProProValLysSerProPro 1077

Qy       1310 CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTC 1369
Db       1078 ProProAlaProValSerSerProProProProValLysSerProPro------------ 1093

Qy       1370 ACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATC 1429
Db       1094 ------------------ProPro---------------AlaProValSerSerPro--- 1101

Qy       1430 CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
Db       1102 ------ProProProIleLysSerProProProProAlaProValSerSerProProPro 1119

Qy       1490 GCT---------CCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAG 1540
Db       1120 AlaProValLysPro-ProSerLeuProProProAlaProValSerSerProPro----- 1137

Qy       1541 CTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCATA 1589
Db       1138 ----ProValValThrProAlaProProLysLysGluGluGlnSerLeu 1152
```

RESULT 10
D86731
hypothetical protein yihD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86731
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrl
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: D86731
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1063 <STO>
A;Cross-references: UNIPROT:Q9CH86; UNIPARC:UPI00000D4407; GB:AE005176; PID:g12723779;
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yihD

Alignment Scores:

Pred. No.:	7.56e-09	Length:	1063
Score:	284.50	Matches:	133
Percent Similarity:	41.1%	Conservative:	50
Best Local Similarity:	29.9%	Mismatches:	191
Query Match:	9.2%	Indels:	71
DB:	2	Gaps:	13

US-10-677-865A-2 (1-1590) x D86731 (1-1063)

```
Qy        88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
Db       594 SerLeuThrProSerValGlyAlaSerThr-------------------------Val 604

Qy       148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
Db       605 AsnProAsnSerSerAlaAsnSerSerGlnSerSerSerSerAlaSerSerSerSerSer 624

Qy       208 GGCTCCAGCAGTTCTGGCAGCTCC------ATTTCCAGTGCCAGAAGCTCTGGTGGTGGC 261
Db       625 AlaAlaSerSerSerThrSerSerSerAsnValSerSerAsnThrSerSerAsnSerSer 644

Qy       262 TCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAG 321
Db       645 GluAlaAsnThrSerSerSerThrSerAsnAlaSerSerSerSerSerSerSerGluGly 664

Qy       322 CCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGT 381
Db       665 SerSerAlaSerSerSerAsnSerSerGluSerSerValAlaSerSerSerSerValAsp 684

Qy       382 GCATCCGGTTCCTCCCAGCTGGGG---------AGCAGCAGCTCTCACTCGGGAAGCAGC 432
Db       685 SerSerGlnSerSerSerAlaGlyValAsnSerSerSerSerSerAlaGluGlySerSer 704

Qy       433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492
Db       705 AlaSerSerSerAsnSerSerGluSerSerValAlaSerSerSerSerSerValAspSerSer 724

Qy       493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552
Db       725 GlnSerSerSerAlaGlyValAsnGlySerSer---SerSerSerGluSerSerSerAla 743

Qy       553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612
Db       744 SerSerSerAsnSerSerGluGlySerValAlaSerSerSerSerValAspSerSerGln 763

Qy       613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672
Db       764 SerSerSerAlaGlyValAsnSerSerSerSerSerAlaGluGlySerSerAlaSerSer 783

Qy       673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
Db       784 SerAsnSerSerGluSerSerIleAlaSerSerSerSer------ValGlySerSerGln 801

Qy       733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792
Db       802 SerSerSerThrGly--------------------ValSerSerSerSerSerSerAla 814

Qy       793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
Db       815 -----------GluGlySerSerAlaSerSerSerAsnSerSerGluSerSerValAla 830

Qy       853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGAC 912
Db       831 SerSerSerSerValAspSerSerGlnSerSerSerAlaGlyValAsnSerSerSerSer 850

Qy       913 AGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTC 972
Db       851 Ser-----------------------SerGluGly------------------------ 854

Qy       973 ACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCC 1032
Db       855 ------------------------------------SerSerAlaSerSerSerAsn 861

Qy      1033 ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCC 1092
Db       862 PheSerGluSerSerIleGlySerSerSerSerValAspSerSerGlnSerSerSerSer 881

Qy      1093 TCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC---GGC 1149
Db       882 ThrGlu-----------------GlySerSerAlaSerSerLeuSerSerSerGluSer 895

Qy      1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
```

```
Db       896 SerMetThrSerSerGlyGlyThrSerAspSerAsnSerAsnThrGlyGlnThrSerThr 915

Qy      1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
Db       916 ThrSerSerSerSerLysAsnSerSerHisSerSerLeuSerSerAsnGluSer------ 933

Qy      1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCG------AGTGGC 1323
Db       934 ---------SerSerAlaSerLeuThrSerSerGluSerSerGlySerThrThrSerSer 950

Qy      1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
Db       951 SerMetLeuThrAsnProAsnLeuGlnSerAsnAspLysAsnGlySerSerLeuThrSer 970

Qy      1384 GTCTCCTCCTTGACA 1398
Db       971 AsnSerSerTyrThr 975
```

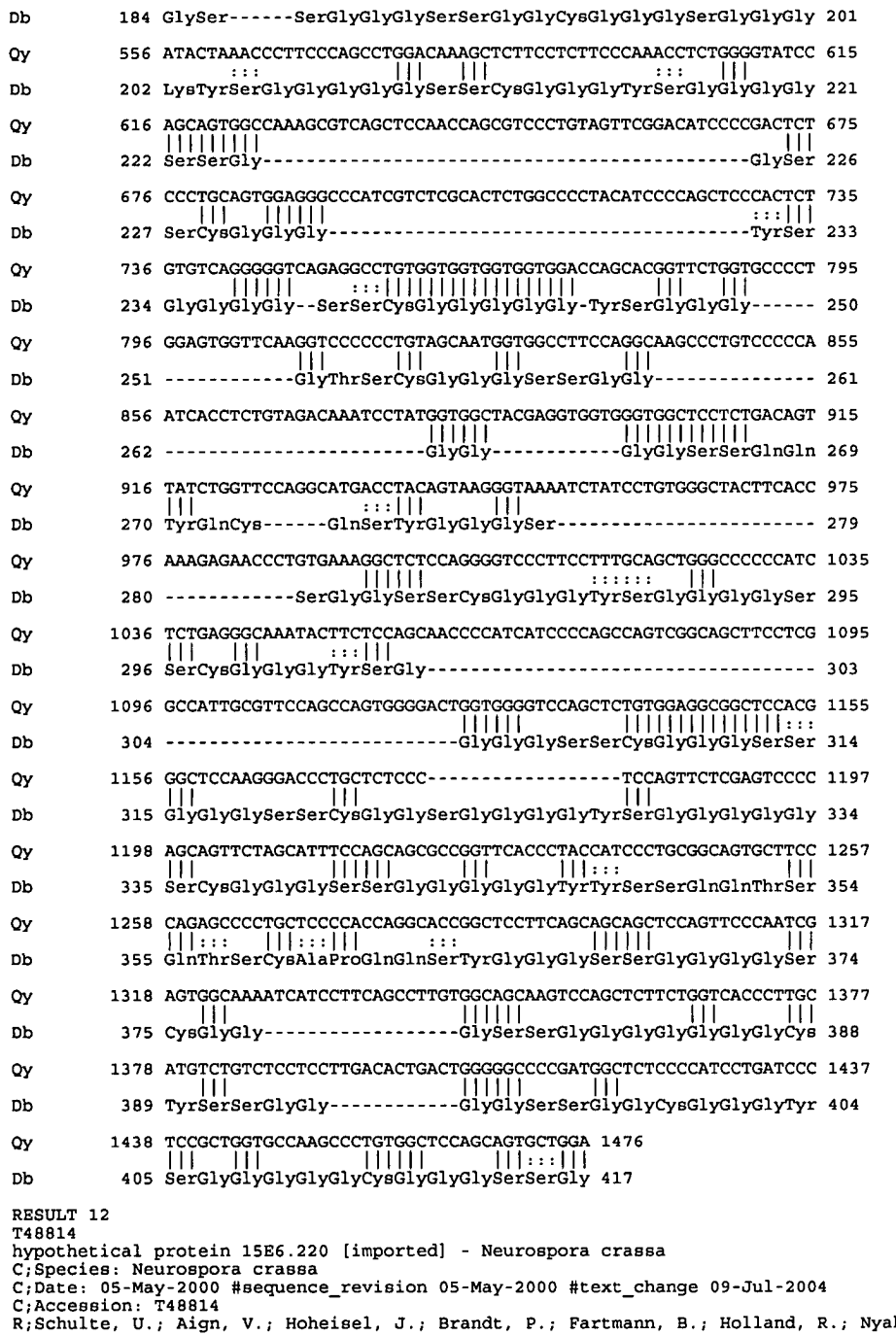
RESULT 11
A35628
loricrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 09-Jul-2004
C;Accession: A35628
R;Mehrel, T.; Hohl, D.; Rothnagel, J.A.; Longley, M.A.; Bundman, D.; Cheng, C.; Lichti,
Cell 61, 1103-1112, 1990
A;Title: Identification of a major keratinocyte cell envelope protein, loricrin.
A;Reference number: A35628; MUID:90275605; PMID:2190691
A;Accession: A35628
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-481 <MEH>
A;Cross-references: UNIPROT:P18165; UNIPARC:UPI00000272CA; GB:M34398; NID:g198870; PIDN
C;Superfamily: loricrin
C;Keywords: cornified cell envelope; epidermis

Alignment Scores:

Pred. No.:	3.56e-08	Length:	481
Score:	273.00	Matches:	127
Percent Similarity:	31.4%	Conservative:	16
Best Local Similarity:	27.9%	Mismatches:	193
Query Match:	8.8%	Indels:	119
DB:	2	Gaps:	16

US-10-677-865A-2 (1-1590) x A35628 (1-481)

```
Qy       172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT---------GGCTCCAGCAGTTCT 222
Db        64 GlyGlyGlyGlySerGlyGlyGlyValLysTyrSerGlyGlyGlyGlyGlySerSerCys 83

Qy       223 GGCAGCTCCATTTCC---------AGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGC 273
Db        84 GlyGlyGlyTyrSerGlyGlyGlyGlyGlySerSerCysGlyGlyGlyTyrSerGlyGly 103

Qy       274 TCCAGCGGATCCAGCATTGCC---------CAGGGTGGTTCTGCAGGATCTTTTAAGCCA 324
Db       104 GlyGlyGlySerSerCysGlyGlyGlyTyrSerGlyGlyGlyGlyGlySerSerCysGly 123

Qy       325 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCA 384
Db       124 GlyGlySerTyrSerGlyGlyGlySerSerCysGlyGlyGlyGlyGlySerGlyGlyGly 143

Qy       385 ---------TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGC 435
Db       144 ValLysTyrSerGlyGlyGlyGlyGlySerSerCysGlyGlyGlySerSerGly 163

Qy       436 TCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGC 495
Db       164 GlyGlyGlyGlyGlySerSerCysGlyGlyGlySerGlyGlyGlyGlySerTyrCysGly 183

Qy       496 AGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGA 555
```



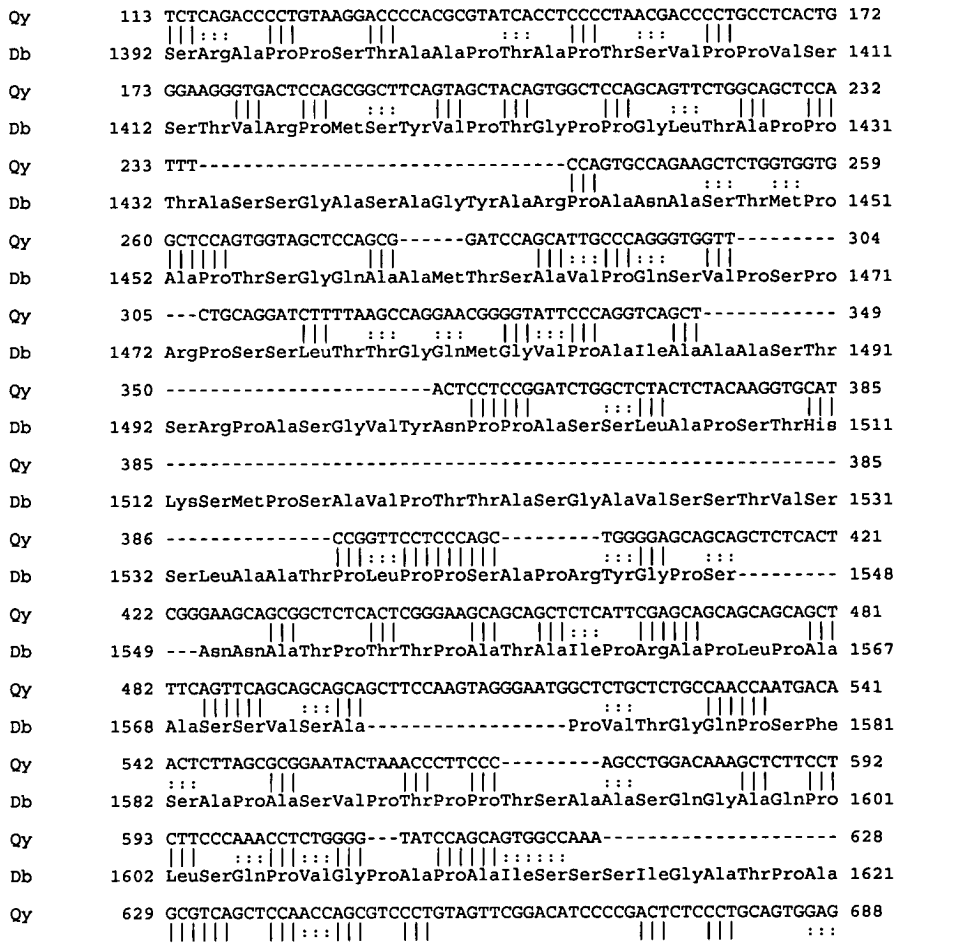
RESULT 12
T48814
hypothetical protein 15E6.220 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: T48814
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48814
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1952 <SCH>
A;Cross-references: UNIPROT:Q9P6T1; UNIPARC:UPI000017B444; EMBL:AL353822; GSPDB:GN00112
A;Experimental source: cosmid contig 15E6; strain 74
C;Genetics:
A;Gene: NCSP:15E6.220
A;Map position: 2
A;Introns: 281/3

Alignment Scores:

Pred. No.:	3.96e-08	Length:	1952
Score:	271.50	Matches:	163
Percent Similarity:	36.4%	Conservative:	59
Best Local Similarity:	26.7%	Mismatches:	218
Query Match:	8.8%	Indels:	170
DB:	2	Gaps:	33

US-10-677-865A-2 (1-1590) x T48814 (1-1952)



```
Db       1622 AlaSerIleProSerSerAlaProAlaAlaLeuAlaProValThrTyrProValProGln 1641

Qy        689 GGCCCATCGTCTCGCACTCTG---GCCCCTACATCCCCAGCTCCC---------ACTCTG 736
              ||:::               |||   :::|||||||||         |||:::
Db       1642 GlnAlaSerAlaAlaAlaAlaArgLeuProValThrProAlaProAlaAlaHisThrIle 1661

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
              |||   |||                                 |||       |||:::
Db       1662 AlaGlnSerVal-----------------------AlaGlnProValAlaArgProVal 1673

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCA------ATGGTGGCCTTCCAGGCAAGC------ 844
                   |||      ||||||      :::   :::   |||||||||
Db       1674 ThrGlnSerProValGlnSerValAlaGlnHisValThrSerSerGlnAlaSerThrThr 1693

Qy        845 ---------CCTGTCCCCCAATCA---------CCTCTGTAGACAAATCCTATGGTGGCT 886
                  ||||||   ||||||          |||:::   :::||||||      |||
Db       1694 AlaAlaHisProValAlaGlnSerValProArgProValAsnSerAsnProThrSerAla 1713

Qy        887 ACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGG 946
                     ||||||:::      :::                        |||   :::
Db       1714 --------------AlaProValAlaAlaVal--------------GlyThrAlaGln 1723

Qy        947 GTAAAATCTATCCTGTGGGCTACT------TCACCAAAGAGAACCCTGTGAAAGGCTCTC 1000
              |||   |||          |||      :::|||   |||   |||      ::::::
Db       1724 ValAlaSerAlaProThrValThrGlnProAlaProHisArgAlaLeu---SerSerVal 1742

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCA 1060
                   |||||||||      :::      |||                   :::
Db       1743 SerGlnSerLeuProGlnSerVal------ProHisAlaAlaGlnGlnAlaAlaHisGln 1760

Qy       1061 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGA 1120
              ||||||   |||   |||||||||   :::|||:::   :::      |||
Db       1761 ThrProHisSer---AlaSerArgProValProGlnSerValProGlnSer--------- 1776

Qy       1121 CTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCT 1180
                 |||      ::::::   :::|||   ||||||   |||         ||||||
Db       1777 ---ValProGlnAlaThrProGlnAlaValProArg---ProSerThrSerAlaLeuThr 1794

Qy       1181 CCAGTTCTCGAG------------TCCCCAGCAGTT------CTAGCATTTCCAGCAGCG 1222
              |||      ::            ||||||||||||                |||||
Db       1795 ProThrAlaGlnProGlyProValSerProAlaValSerGlySerGlyValProAlaPro 1814

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
                    ::::::      |||   :::         |||   |||
Db       1815 SerAlaAlaGlnSerValAlaProAlaProValSerSerThrProValProAlaAlaThr 1834

Qy       1283 CCGGCTCCTTCA---------GCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCC 1333
                 ||||||:::         ||||||||||||   |||   ||||||      :::
Db       1835 ValAlaProAlaSerThrValAlaAlaAlaProThrProThrArgValThrAlaAlaPro 1854

Qy       1334 TTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCT 1393
                 :::|||   ||||||:::            |||                |||   |||
Db       1855 AlaAlaLeuSerAlaAlaThrAsnProAlaProValPro------------SerGlnPro 1870

Qy       1394 TGACACTGACTG---------------------------------------------- 1405
                 |||   :::
Db       1871 GlnHisGlnIleThrGlyGlnAlaProAlaGlnGlnGlnArgProProAlaGlnAlaGln 1890

Qy       1406 --------------------------GGGGCCCCGATGGCTCTCCCCATCCTGATCCCT 1438
                                       ||||||          |||   |||   |||
Db       1891 AlaProAlaThrProThrIleThrSerAlaAlaProProArgProProThrLeuAlaPro 1910

Qy       1439 CCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC 1498
              |||      |||   |||      |||          ||| ||||||   |||      ||
Db       1911 Pro------ProProProProProProProProThrGlu-AspProProPro------Pr 1926

Qy       1499 GGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCC 1558
              |    |||   |||:::||||||          |||                    ||
Db       1926 oProProProProProAlaGluAlaProProProProPro------------------Pr 1940

Qy       1559 TACCCCAAGGAGAGTTACTCGACAGTCC 1586
              |||||||         |||      |||
Db       1940 oThrProLeuMetProThrSerAlaSer 1949
```

```
RESULT 13
A47282
calcium-binding protein calphotin - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47282
R;Martin, J.H.; Benzer, S.; Rudnicka, M.; Miller, C.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 1531-1535, 1993
A;Title: Calphotin: a Drosophila photoreceptor cell calcium-binding protein.
A;Reference number: A47282; MUID:93165729; PMID:8094559
A;Accession: A47282
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-865 <MAR>
A;Cross-references: UNIPROT:Q02910; UNIPARC:UPI00001282D3; GB:L02111; NID:g157031; PIDN
A;Experimental source: photoreceptor cells
A;Note: sequence extracted from NCBI backbone (NCBIN:124955, NCBIP:124956)
C;Genetics:
A;Gene: FlyBase:Cpn
A;Cross-references: FlyBase:FBgn0010218
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology
C;Keywords: calcium binding

Alignment Scores:
Pred. No.:                4.45e-08      Length:          865
Score:                    271.00        Matches:         143
Percent Similarity:       35.5%         Conservative:    70
Best Local Similarity:    23.8%         Mismatches:      194
Query Match:              8.7%          Indels:          193
DB:                       2             Gaps:            26

US-10-677-865A-2 (1-1590) x A47282 (1-865)

Qy        101 GCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACC 160
              ::::::||||||                    ||||||      :::   |||::::::|||
Db         80 SerValAlaProProValValAlaAlaProThrProProAlaAlaSerProValSerThr 99

Qy        161 CCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTT 220
              |||   ::::::   ::::::   |||   |||       |||   |||||||||   |||
Db        100 ProValAlaValAlaGlnIle---ProValAlaValSerAlaProValAlaProProVal 118

Qy        221 CTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCT------ 274
                 |||   |||   ||||||   :::       |||   |||||||||:::|||
Db        119 AlaAlaThrProThrProValValGlnIleProValAlaAlaProValIleAlaThrPro 138

Qy        275 ---------------CCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTA 319
                            ||||||                ||:::      ::::::
Db        139 ProValAlaAlaSerAlaProThrProAlaAlaValThrProValIleSerProValIle 158

Qy        320 AGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAG 379
              :::            :::|||   :::   |||   |||
Db        159 AlaSerProProValValProAlaAsnThrThrValPro--------------------- 171

Qy        380 GTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTC 439
                                      ||||||   :::          ||||||:::
Db        172 ----------------------ValAlaAlaProValAlaAlaValProAlaAlaVal 183

Qy        440 ACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
                     |||      |||      |||   |||   |||   :::         |||
Db        184 ProValValAlaProValLeuAlaProAlaValAlaProAlaValAlaProValValAla 203

Qy        500 GCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATAC 559
                 :::
Db        204 GluThr---------------------------------------------------- 205

Qy        560 TAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGG--------- 610
                                      |||   |||   |||:::
Db        206 ----------------------------ProAlaProProProValAlaGluIlePro 215

Qy        611 ----------------------TATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGC 646
                                   |||   :::   :::   |||         :::
Db        216 ValAlaThrIleProGluCysValAlaProLeuIleProGluValSerValValAlaThr 235
```

```
Qy       647 GTCCCTGTAGTTCGGACATCCCCGACTCTC---------CCTGCAGTGGAGGGCCCATCG 697
Db       236 LysProLeuAlaAlaAlaGluProValValValAlaProProAlaThrGluThrPro--- 254

Qy       698 TCTCGCACTCTGGCCCCTACATCCCCAGCTCCC------ACTCTGTGTCAGGGGGTCAGA 751
Db       255 ------ValValAlaProAlaAlaAlaSerProHisValSerValAlaProAlaValGlu 272

Qy       752 GGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTC 811
Db       273 ----------------------ThrAlaValValAlaProValSerAlaSerThrGlu 284

Qy       812 CCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCT---------------------- 847
Db       285 ProProValAlaAlaAlaThrLeuThrThrAlaProGluThrProAlaLeuAlaProVal 304

Qy       848 GTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACG----------------- 889
Db       305 ValAlaGluSerGlnValAlaAlaAsnThrValValAlaThrProProThrProAlaPro 324

Qy       890 ---AGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGG 946
Db       325 GluProGluThrIleAlaProProValValAla--------------GluThrProGlu 339

Qy       947 GTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGG 1006
Db       340 ValAlaSerValAlaValAlaGluThrThrProProValValProProValAlaAlaGlu 359

Qy      1007 TCCCTTCCT-----------------------------------TTGCAGCTGGGCCCC 1030
Db       360 SerIleProAlaProValValAlaThrThrProValProAlaThrLeuAlaValThrAsp 379

Qy      1031 CCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTT 1090
Db       380 ProAspValThrAlaSerAlaValProGluLeuProProValIleAla------------ 395

Qy      1091 CCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCT 1150
Db       396 ProSerProValProSerAla------------ValAlaGluThrProValAspLeuAla 411

Qy      1151 CCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTC---------------------- 1189
Db       412 Pro---------------ProValLeuProProValAlaAlaGluProValProAlaVal 426

Qy      1190 ---------GAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATC 1240
Db       427 ValAlaGluGluThrProGluThrProAlaProAlaSerAlaProVal------ThrIle 444

Qy      1241 CCTGCGGCAGTGCTTCCCAGA---GCCCCT---------GCTCCCCACCAGGCACCG--- 1285
Db       445 AlaAlaLeuAspIleProGluValAlaProValIleAlaAlaProSerAspAlaProAla 464

Qy      1286 ---GCTCCTTCAGCAGCAGCTCCAGTT------------------------------- 1309
Db       465 GluAlaProSerAlaAlaAlaProIleValSerThrProProThrThrAlaSerValPro 484

Qy      1310 --------------CCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGT 1354
Db       485 GluThrThrAlaProProAlaAlaValProThrGluProIleAspValSerValLeuSer 504

Qy      1355 CCAGCTCTTCTG----------------------------------------GTCACCCTT 1375
Db       505 GluAlaAlaIleGluThrProValAlaProProValGluValThrThrGluValAlaVal 524

Qy      1376 GCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATC 1435
Db       525 AlaAspValAlaProPro-----------------GluAlaAlaAlaAspLeuIleIle 538

Qy      1436 CCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGA-------------- 1480
Db       539 GluProValGluProProAlaProIleProAspLeuLeuGluGlnThrThrSerValPro 558

Qy      1481 --------------------------TCCCCTGCCGCTCCATCCGGGATATCCTAGCCC 1513
Db       559 AlaValGluAlaAlaGluSerThrSerSerProIleProGluThrSerLeu-ProProPr 578

Qy      1514 AAGTGAAGCCTCTGGGGCCCCA------GCTAGCTGACCCTGAAGTTTTCCTACCCCA 1565
Db       578 oAsnGluAlaValAlaSerProGluValAlaValAlaProIleThrAlaProGluPro 597
```

RESULT 14
T33369
hypothetical protein H02F09.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33369
R;Geisel, C.; Harmon, G.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid H02F09.
A;Reference number: Z21330
A;Accession: T33369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1275 <GEI>
A;Cross-references: UNIPROT:O76602; UNIPARC:UPI0000082AAD; EMBL:AF077538; PIDN:AAC64622
A;Experimental source: strain Bristol N2; clone H02F09
C;Genetics:
A;Gene: CESP:H02F09.3
A;Map position: X
A;Introns: 42/1; 78/3; 106/3; 135/2; 161/2; 224/1
C;Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase

Alignment Scores:

Pred. No.:	4.34e-08	Length:	1275
Score:	271.00	Matches:	136
Percent Similarity:	41.9%	Conservative:	79
Best Local Similarity:	26.5%	Mismatches:	226
Query Match:	8.7%	Indels:	72
DB:	2	Gaps:	22

US-10-677-865A-2 (1-1590) x T33369 (1-1275)

```
Qy        79 CTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACG 138
Db       659 IleProGlySerThrGluSerThrVal--------------------SerGluAlaSer 671

Qy       139 CGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG------AAGGGTGACTCCAGCGGC 192
Db       672 ThrValSerGlySerSerValSerThrValSerGlySerThrGluSerThrSerAlaGly 691

Qy       193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
Db       692 AlaSerThrValSerGlySerThrGlySerThrValSerAspSerSerThrIleSerAsp 711

Qy       253 GGTGGTGGCTCCAGT---------GGTAGCTCC------------AGCGGATCCAGCATT 291
Db       712 SerThrGlySerThrAsnAlaProGlySerThrGluSerThrValThrGlySerSerVal 731

Qy       292 GCCCAG---GGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
Db       732 SerThrValSerGlySerThrGlySer------------ThrGlyProSerThrMetSer 747

Qy       349 TACTCCTCCGGATCT-------------------------------------GGCTCT 369
Db       748 AlaSerThrGlySerThrAsnThrProGlySerThrGluSerThrIleThrAspGlySer 767

Qy       370 ACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGC 429
Db       768 ThrValSerGlySerThrGlySerThr------GlySerThrAsnAsnProGlySerThr 785

Qy       430 AGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTC 489
Db       786 AspSerSerThrThrGlyIleSerThrValSerGlySerSerLeuSerThrIleSerGly 805

Qy       490 AGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCT---GCTCTGCCAACCAATGACAACTCT 546
```

```
Db        806 SerThrGlySerThrValSerGlySerSerAspMetThrValSerThrGlySerThrSer 825

Qy        547 TAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCT 606

Db        826 SerProGlySerThrGluSerThrValSerGlyAlaSerThrMetSerProSerThrGly 845

Qy        607 GGG---------GTATCCAGCAGTGGCCAAAGCGTCAGCTCC---AACCAGCGTCCCTGT 654

Db        846 SerSerValGluThrSerThrSerGlySerSerValSerThrValSerGlnSerThrSer 865

Qy        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATC------GTCTCGCACTCT 708

Db        866 SerSerThrThrGlyGlnSerThrValSerGluSerSerValSerThrValSerSerGlu 885

Qy        709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768

Db        886 SerThrIleSerGlnSerThrGlySerThrThrThrGlyGluSer---------ThrVal 902

Qy        769 GTGGACCAGCACGGTTCTGGTGCCCCTGGA---GTGGTTCAAGGTCCCCCCTGTAGCAAT 825

Db        903 PheGlySerThrGlySerThrAlaThrGlySerSerThrMetSerAlaSerThrGlySer 922

Qy        826 GGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGC 885

Db        923 ThrAspThrProGlySerThrGluSerThrIleThrGly------SerThrValThrGly 940

Qy        886 TACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAG 945

Db        941 GluSerThrValSerGlySerThrGlySerThrIleThrGluGlySerThrIleSerGlu 960

Qy        946 GGTAAAATCTATCCTGTGGGCTAC--------------TTCACCAAAGAGAACCCTGTG 990

Db        961 SerThrMetThrThrValGlyValSerThrGlySerThrIleThrGlyGluSerThrVal 980

Qy        991 AAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATAC 1050

Db        981 SerGlySer------ThrArgSerThrValThrGlyGluSerThrValSerGlySerThr 998

Qy       1051 TTCTCC--------------------AGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089

Db        999 GluSerThrValSerGlySerThrGluSerThrProThrValProSerThrValSerGly 1018

Qy       1090 TCC------TCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143

Db       1019 SerThrGlySerThrValThrGlyGluSerThrValSerGlySerThrAlaSerThrSer 1038

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203

Db       1039 SerGlySerThrGlySerSerThrGluAlaGlySerThrValSerGly---SerSerAla 1057

Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263

Db       1058 SerThrValThrSerSerThrGlySer--------------SerThrSerGlyGluSer 1072

Qy       1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323

Db       1073 ThrValSerGlySerThrValSerThrValSerGlySerThrGlySerThrIleThrGly 1092

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383

Db       1093 GluSerThrValSerGly---SerThrGluSerThrValThrAlaGluSerThrValSer 1111

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443

Db       1112 GlySerSerValSerThrValSerGlyAsnThrGlySerThrIleThrGlyGluSerThr 1131

Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATC 1482

Db       1132 ValSerGlySerThrGlySerThrGlyGluSerThrIle 1144

RESULT 15
A47283
```

```
calphotin - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: A47283
R;Ballinger, D.G.; Xue, N.; Harshman, K.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 1536-1540, 1993
A;Title: A Drosophila photoreceptor cell-specific protein, calphotin, binds calcium and
A;Reference number: A47283; MUID:93165730; PMID:8434015
A;Contents: photoreceptor cells
A;Accession: A47283
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-873 <BAL>
A;Cross-references: UNIPROT:Q02910; UNIPARC:UPI000016BB46; GB:L05080; NID:g157071; PIDN
A;Note: sequence extracted from NCBI backbone (NCBIN:124958, NCBIP:124959)
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

Alignment Scores:
Pred. No.:              9.11e-08        Length:         873
Score:                  265.50          Matches:        153
Percent Similarity:     35.6%           Conservative:   73
Best Local Similarity:  24.1%           Mismatches:     210
Query Match:            8.6%            Indels:         198
DB:                     2               Gaps:           28

US-10-677-865A-2 (1-1590) x A47283 (1-873)

Qy         95 CTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTA 154

Db         27 ValGlnValValSerProAlaAlaValAlaValAlaProAlaValValAlaProAlaPro 46

Qy        155 ACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCA 214

Db         47 AlaAlaProThrAlaValThrProValAlaProProProThrLeuAlaSerValGlnPro 66

Qy        215 GCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTG------------------ 256

Db         67 AlaThrValThrValProAlaProAlaProIleAlaAlaAlaSerValAlaProValAla 86

Qy        257 ---GTGGCTCCA------GTGGTAGCTCCAGCGGATCCAGCA--------------TTG 292

Db         87 SerValAlaProProValValAlaAlaProThrProProAlaAlaSerProValSerThr 106

Qy        293 CCCAGGGTGGTTCTGCAGGATCTT---------TTAAGCCAGGAACGGGGTATTCCCAGG 343

Db        107 ProProValAlaValAlaGlnIleProValAlaValSerAlaProValAlaProProVal 126

Qy        344 TCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGG 403

Db        127 AlaAlaThrProThrProValAlaProIleProValAlaAlaProValIleAlaThrPro 146

Qy        404 GGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATT 463

Db        147 ProValAlaAlaSerAlaProThrProAlaAlaValThrProValValSerProValIle 166

Qy        464 CGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTG 523

Db        167 AlaThrProProValValProAlaAsnThrThrValPro---------------------- 179

Qy        524 CTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAA 583

Db        180 ValAlaAlaProValAlaAlaValProAla------------------------------ 189

Qy        584 GCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACC 643

Db        190 AlaValProValValAlaProValLeuAlaProAlaValAlaProAlaValAlaProVal 209

Qy        643 ------------------------------------------------------------ 643

Db        210 ValAlaGluThrProAlaProProProValAlaGluIleProValAlaThrIleProGlu 229

Qy        644 ---------------------AGCGTCCCTGTAGTTCGGACATCCCCGACTCTC------ 676
```

```
Db        230 CysValAlaProLeuIleProGluValSerValValAlaThrLysProLeuAlaAlaAla 249

Qy        677 --------------------CCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCT 715
                                  ||||||   |||   |||           :::||||||
Db        250 GluProValValValAlaProProAlaThrGluThrPro---------ValValAlaPro 266

Qy        716 ACATCCCCAGCTCCC------ACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGG 769
                 :::   :::|||      ::::::         |||
Db        267 AlaAlaAlaSerProHisValSerValAlaProAlaValGlu------------------ 280

Qy        770 TGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTG 829
                 :::   |||:::   |||:::                |||||||||||
Db        281 ------ThrAlaValValAlaProValSerAlaSerThrGluProProValAlaAlaAla 298

Qy        830 GCCTTCCAGGCAAGCCCT----------------------GTCCCCCAATCACCTCTG 865
                     :::|||                         |||   :::|||   :::
Db        299 ThrLeuThrThrAlaProGluThrProAlaLeuAlaProValValAlaGluSerGlnVal 318

Qy        866 TAGACAAATCCTATGGTGGCTACG--------------------AGGTGGTGGGTGGCT 904
                  |||   :::|||||||||                        :::|||
Db        319 AlaAlaAsnThrValValAlaThrProProThrProAlaProGluProGluThrIleAla 338

Qy        905 CCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGG 964
                 |||      |||                |||     |||   |||:::
Db        339 ProProValValAla---------------GluThrProGluValAlaSerValAlaVal 353

Qy        965 GCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCT--------- 1015
                 |||   :::           :::                  |||:::|||
Db        354 AlaGluThrThrProProValValProProValAlaAlaGluSerIleProAlaProVal 373

Qy       1016 --------------------------TTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 1048
                                        |||   :::      |||   :::   |||:::
Db        374 ValAlaThrThrProValProAlaThrLeuAlaValThrAspProAspValThrAlaSer 393

Qy       1049 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 1108
                  |||      |||          |||              |||   |||:::   |||
Db        394 AlaValProGluLeuProProValIleAla-----------ProSerProValProSer 409

Qy       1109 AGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGAC 1168
                 :::            |||      :::   |||:::   ||||||
Db        410 Ala-----------ValAlaGluThrProValAspLeuAlaPro--------------- 420

Qy       1169 CCTGCTCTCCCTCCAGTTCTC----------------------------GAGTCCCCA 1198
                 |||   ||||||||||||                            |||:::|||
Db        421 ProValLeuProProValAlaAlaGluProValProAlaValValAlaGluGluThrPro 440

Qy       1199 GCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCC 1258
                    |||      :::|||||||||      ||||||   |||      :::|||
Db        441 GluThrProAlaProAlaSerAlaProVal------ThrIleAlaAlaLeuAspIlePro 458

Qy       1259 AGA---GCCCCT---------GCTCCCCACCAGGCACCG------GCTCCTTCAGCAGCA 1300
                 ||||||         ||||||      ||||||      |||||||||||||||
Db        459 GluValAlaProValIleAlaAlaProSerAspAlaProAlaGluAlaProSerAlaAla 478

Qy       1301 GCTCCAGTT-------------------------------------------------CCC 1312
                 ||||||:::                                             |||
Db        479 AlaProIleValSerThrProProThrThrAlaSerValProGluThrThrAlaProPro 498

Qy       1313 AATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTG------ 1366
                    |||          :::          |||   |||   :::
Db        499 AlaAlaValProThrGluProIleAspValSerValLeuSerGluAlaAlaIleGluThr 518

Qy       1367 --------------------------------GTCACCCTTGCATGTCTGTCTCCTCCT 1393
                                             |||   :::|||   ::::::||||||
Db        519 ProValAlaProProValGluValThrThrGluValAlaValAlaAspValAlaProPro 538

Qy       1394 TGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGC 1453
                                        |||      ::::::|||   |||:::   |||
Db        539 -----------------GluAlaAlaGluAspLeuIleIleGluProValGluProPro 552

Qy       1454 CCTGTGGCTCCAGCAGTGCTGGAAAGA------------------------------- 1480
                    |||   :::||||||:::
Db        553 AlaProIleProAspLeuLeuGluGlnThrThrSerValProAlaValGluAlaAlaGlu 572

Qy       1481 ---------TCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGC 1531
                         ||||||         :::   ::: |||   |||:::||||||      ::
Db        573 SerThrSerSerProIleProGluThrSerLeu-ProProProAsnGluAlaValAlaSe 592

Qy       1532 CCCA------GCTAGCTGACCCTGAAGTTTTCCTACCCCA 1565
                  :|||      |||      |||   :::   |||   |||
Db        592 rProGluValAlaValAlaProIleThrAlaProGluPro 605
```

Search completed: March 21, 2006, 22:07:27
Job time : 78.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 21:53:34 ; Search time 42 Seconds
(without alignments)
5341.857 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152017_26068/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_192_@runat_21032006_152017_26068 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2735	88.2	529	1	CDSN_HUMAN	Q15517 homo sapien
2	2717	87.6	529	1	CDSN_PANTR	Q7yr44 pan troglod
3	2492.5	80.4	534	1	CDSN_MACMU	Q5tm45 macaca mula
4	1746	56.3	561	1	CDSN_MOUSE	Q7tpc1 mus musculu
5	974	31.4	225	1	CDSN_PIG	O19084 sus scrofa
6	827	26.7	262	2	Q767M8_PIG	Q767m8 sus scrofa
7	693	22.3	133	2	Q7Z559_HUMAN	Q7z559 homo sapien
8	692	22.3	133	2	Q7Z560_HUMAN	Q7z560 homo sapien
9	392	12.6	2425	2	Q4FX61_LEIMA	Q4fx61 leishmania
10	387.5	12.5	1435	2	Q4FX64_LEIMA	Q4fx64 leishmania
11	387.5	12.5	17392	2	Q4FX62_LEIMA	Q4fx62 leishmania
12	382	12.3	7194	2	Q4FX63_LEIMA	Q4fx63 leishmania
13	375.5	12.1	1325	2	Q9BKV7_LEIMA	Q9bkv7 leishmania
14	351	11.3	669	2	Q7UE67_RHOBA	Q7ue67 rhodopirell
15	341	11.0	374	2	Q75JC9_DICDI	Q75jc9 dictyosteli
16	341	11.0	825	2	Q7YZW4_CAEEL	Q7yzw4 caenorhabdi
17	341	11.0	1032	2	P91365_CAEEL	P91365 caenorhabdi
18	341	11.0	2232	2	Q8IFX6_CAEEL	Q8ifx6 caenorhabdi
19	337	10.9	13288	2	O18758_PIG	O18758 sus scrofa
20	335.5	10.8	950	2	Q9VE45_DROME	Q9ve45 drosophila
21	335	10.8	1416	2	Q6VBJ3_CANGA	Q6vbj3 candida gla
22	330.5	10.7	1218	2	Q6VBJ4_CANGA	Q6vbj4 candida gla
23	325.5	10.5	774	2	Q868B4_CAEEL	Q868b4 caenorhabdi
24	325	10.5	1035	2	Q76C74_YEAST	Q76c74 saccharomyc
25	325	10.5	1713	2	Q8TGE1_YEAST	Q8tge1 saccharomyc
26	323.5	10.4	1428	2	O44341_HALRU	O44341 haliotis ru
27	318.5	10.3	4202	2	Q4IDY5_GIBZE	Q4idy5 gibberella
28	318	10.3	1514	2	Q4FX85_LEIMA	Q4fx85 leishmania
29	315.5	10.2	1553	2	Q6FPR0_CANGA	Q6fpr0 candida gla
30	315	10.2	2112	2	Q9VEL9_DROME	Q9vel9 drosophila
31	312.5	10.1	1376	2	Q7S5H8_NEUCR	Q7s5h8 neurospora
32	312	10.1	486	2	Q54KD5_DICDI	Q54kd5 dictyosteli
33	309.5	10.0	2850	1	HORN_HUMAN	Q86yz3 homo sapien
34	309.5	10.0	2850	2	Q5U1F4_HUMAN	Q5u1f4 homo sapien
35	309.5	10.0	2850	2	Q5DT20_HUMAN	Q5dt20 homo sapien
36	308.5	9.9	967	2	Q08294_YEAST	Q08294 saccharomyc
37	306	9.9	1150	1	MUCAP_PIG	P12021 sus scrofa
38	305.5	9.8	1143	2	Q54GV8_DICDI	Q54gv8 dictyosteli
39	301	9.7	1198	2	Q7PKP4_ANOGA	Q7pkp4 anopheles g
40	301	9.7	1815	2	Q6CF66_YARLI	Q6cf66 yarrowia li
41	300	9.7	646	2	Q61QV1_CAEBR	Q61qv1 caenorhabdi
42	299.5	9.7	763	2	Q9XDH2_MYCTU	Q9xdh2 mycobacteri
43	299.5	9.7	784	2	Q55F46_DICDI	Q55f46 dictyosteli
44	297	9.6	766	2	Q6CV63_KLULA	Q6cv63 kluyveromyc
45	293	9.4	2532	2	Q629H6_CAEBR	Q629h6 caenorhabdi

ALIGNMENTS

RESULT 1
CDSN_HUMAN
ID CDSN_HUMAN STANDARD; PRT; 529 AA.
AC Q15517; O43509; Q5SQ85; Q5STD2; Q7LA70; Q7LA71; Q86Z04; Q8IZU4;
AC Q8IZU5; Q8IZU6; Q8N5P3; Q95IF9; Q9NP52; Q9NPE0; Q9NPG5; Q9NRH4;
AC Q9NRH5; Q9NRH6; Q9NRH7; Q9NRH8; Q9UBH8; Q9UIN6; Q9UIN7; Q9UIN8;
AC Q9UIN9; Q9UIP0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Corneodesmosin precursor (S protein).
GN Name=CDSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LEU-18.
RC TISSUE=Foreskin;
RX MEDLINE=94022396; PubMed=8415725;
RA Zhou Y., Chaplin D.D.;
RT "Identification in the HLA class I region of a gene expressed late in
RT keratinocyte differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9470-9474(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT ALA-408.
RC TISSUE=Epidermis;
RX MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT "Expression cloning of human corneodesmosin proves its identity with
RT the product of the S gene and allows improved characterization of its
RT processing during keratinocyte differentiation.";
RL J. Biol. Chem. 273:22640-22647(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS LEU-18; SER-153 DEL;
RP PHE-202 AND LEU-410.
RX MEDLINE=22256218; PubMed=12366786;
RX DOI=10.1034/j.1399-0039.2002.600110.x;
RA Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,

RA Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT patients with psoriasis.";
RL Tissue Antigens 60:77-83(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RT "Genome diversity in HLA: a new strategy for detection of genetic
RT polymorphisms in expressed genes within the HLA class III and class I
RT regions.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANTS LEU-18;
RP ASN-143; SER-153 DEL; PHE-202; ALA-408; LEU-410 AND ASN-527.
RX MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcott R.,
RA Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLay K.,
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT "The DNA sequence and analysis of human chromosome 6.";
RL Nature 425:805-811(2003).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS LEU-18; SER-153
RP DEL; PHE-202 AND LEU-410.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 30-529 (ALLELES 1.11; 1.21; 1.31;
RP 1.32; 1.41; 1.42; 1.43; 1.51; 1.52; 2.11; 2.21; 2.22 AND 2.23), AND
RP VARIANTS PHE-56; ASN-143 DEL; ASN-150; PHE-202; GLY-401; ALA-408;
RP LEU-410 AND ASN-527.
RX MEDLINE=21108612; PubMed=11169256;
RX DOI=10.1034/j.1399-0039.2001.057001032.x;
RA Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT "Identification of six novel polymorphisms in the human corneodesmosin
RT gene.";
RL Tissue Antigens 57:32-38(2001).
RN [9]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 106-529, VARIANTS ASN-143;
RP PHE-202; GLY-401; ALA-408 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX MEDLINE=20066866; PubMed=10599883;
RX DOI=10.1034/j.1399-0039.1999.540501.x;
RA Jenisch S., Koch S., Henseler T., Nair R.P., Elder J.T., Watts C.E.,
RA Westphal E., Voorhees J.J., Christophers E., Kroenke M.;
RT "Corneodesmosin gene polymorphism demonstrates strong linkage
RT disequilibrium with HLA and association with psoriasis vulgaris.";
RL Tissue Antigens 54:439-449(1999).
RN [10]
RP PARTIAL PROTEIN SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=98058975; PubMed=9395522; DOI=10.1074/jbc.272.50.31770;
RA Simon M., Montezin M., Guerrin M., Durieux J.-J., Serre G.;
RT "Characterization and purification of human corneodesmosin, an
RT epidermal basic glycoprotein associated with corneocyte-specific
RT modified desmosomes.";
RL J. Biol. Chem. 272:31770-31776(1997).
RN [11]
RP INVOLVEMENT IN HTSS.
RX PubMed=12754508; DOI=10.1038/ng1163;
RA Levy-Nissenbaum E., Betz R.C., Frydman M., Simon M., Lahat H.,
RA Bakhan T., Goldman B., Bygum A., Pierick M., Hillmer A.M., Jonca N.,
RA Toribio J., Kruse R., Dewald G., Cichon S., Kubisch C., Guerrin M.,
RA Serre G., Nothen M.M., Pras E.;
RT "Hypotrichosis simplex of the scalp is associated with nonsense
RT mutations in CDSN encoding corneodesmosin.";
RL Nat. Genet. 34:151-153(2003).
RN [12]
RP VARIANTS PHE-202 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX PubMed=10844560; DOI=10.1046/j.1523-1747.2000.00002.x;
RA Enerbaeck C., Enlund F., Inerot A., Samuelsson L., Wahlstroem J.,
RA Swanbeck G., Martinsson T.;
RT "S gene (Corneodesmosin) diversity and its relationship to psoriasis;
RT high content of cSNP in the HLA-linked S gene.";
RL J. Invest. Dermatol. 114:1158-1163(2000).
RN [13]
RP VARIANTS PHE-202; GLY-401; ALA-408; LEU-410 AND ASN-527, AND
RP ASSOCIATION WITH PSORIASIS.
RX PubMed=12472658; DOI=10.1034/j.1399-0039.2002.600403.x;
RA Orru S., Giuressi E., Casula M., Loizedda A., Murru R., Mulargia M.,
RA Masala M.V., Cerimele D., Zucca M., Aste N., Biggio P., Carcassi C.,
RA Contu L.;
RT "Psoriasis is associated with a SNP haplotype of the corneodesmosin
RT gene (CDSN).";
RL Tissue Antigens 60:292-298(2002).
CC -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC intercellular structures that are involved in desquamation.
CC -!- TISSUE SPECIFICITY: Exclusively expressed in skin.
CC -!- POLYMORPHISM: Genetic variation in CDSN may be associated with
CC susceptibility to psoriasis [MIM:177900]. Various CDSN alleles are
CC known including alleles 1.11, 1.21, 1.31, 1.32, 1.41, 1.42, 1.43,
CC 1.51, 1.52, 2.11, 2.21, 2.22 and 2.23.
CC -!- DISEASE: Defects in CDSN are a cause of hypotrichosis simplex of
CC the scalp (HTSS) [MIM:146520]; also known as hypotrichosis Spanish

```
CC       type. HTSS is an autosomal dominant form of isolated alopecia.
CC       Affected individuals have normal hair in early childhood but
CC       experience progressive loss of scalp hair beginning in the middle
CC       of the first decade and almost complete baldness by the third
CC       decade.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 501.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L20815; AAA21321.1; ALT_SEQ; mRNA.
DR   EMBL; AF030130; AAC24196.1; -; mRNA.

Alignment Scores:
Pred. No.:              4.56e-138     Length:          529
Score:                  2735.00       Matches:         526
Percent Similarity:     99.8%         Conservative:    2
Best Local Similarity:  99.4%         Mismatches:      1
Query Match:            88.2%         Indels:          0
DB:                     1             Gaps:            0

US-10-677-865A-2 (1-1590) x CDSN_HUMAN (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529

RESULT 2
CDSN_PANTR
ID   CDSN_PANTR     STANDARD;      PRT;   529 AA.
AC   Q7YR44;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; BA000041; BAC78170.1; -; Genomic_DNA.
KW   Glycoprotein; Signal.
FT   SIGNAL        1    32       Potential.
FT   CHAIN        33   529       Corneodesmosin.
FT   COMPBIAS     62   464       Ser-rich.
FT   CARBOHYD    172   172       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;

Alignment Scores:
Pred. No.:              4.17e-137      Length:         529
Score:                  2717.00        Matches:        523
Percent Similarity:     99.2%          Conservative:   2
Best Local Similarity:  98.9%          Mismatches:     4
Query Match:            87.6%          Indels:         0
DB:                     1              Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_PANTR (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProMetArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrPheGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||:::|||   ||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerSerGlyLeuProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520
```

```
Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529

RESULT 3
CDSN_MACMU
ID   CDSN_MACMU     STANDARD;      PRT;   534 AA.
AC   Q5TM45;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=CDSN;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15269276; DOI=10.1093/molbev/msh216;
RA   Kulski J.K., Anzai T., Shiina T., Inoko H.;
RT   "Rhesus macaque class I duplicon structures, organization, and
RT   evolution within the alpha block of the major histocompatibility
RT   complex.";
RL   Mol. Biol. Evol. 21:2079-2091(2004).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AB128049; BAD69749.1; -; Genomic_DNA.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    534       Corneodesmosin.
FT   COMPBIAS     62    469       Ser-rich.
SQ   SEQUENCE   534 AA;  51983 MW;  CB6A9DF256C32874 CRC64;

Alignment Scores:
Pred. No.:              4.09e-125      Length:         534
Score:                  2492.50        Matches:        486
Percent Similarity:     91.7%          Conservative:   12
Best Local Similarity:  89.5%          Mismatches:     22
Query Match:            80.4%          Indels:         23
DB:                     1              Gaps:           3

US-10-677-865A-2 (1-1590) x CDSN_MACMU (1-534)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||   ||||||||||||||||||||||||||||||||   ||||||||||||||||
Db          1 MetGlyLeuSerArgAlaProTrpMetGlyArgValGlyGlyArgGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db         61 GlySerSerGlyPheSerSerTyrSerGlySerSerGlySerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              |||     ||||||||||||          ||||||||||||||:::|||||||||
Db         81 AlaSerGlySerGlyGlyGly------------SerSerGlySerSerValAlaGlnGly 96

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 116

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTG------------------ 402
              |||||||||:::|||||||||||||||   ||||||||||
Db        117 SerGlySerSerLeuGlnGlyAlaSerSerSerSerGlnLeuGlyGlySerGlySerGln 136

Qy        403 ----------------------GGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCT 438
                                     ||||||||   ||||||||||||||||||||||||
Db        137 ProGlySerSerGlySerGlnProGlySerSerGlySerHisSerGlySerSerGlySer 156

Qy        439 CACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGC 498
              ::::::||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        157 AsnThrGlySerSerSerSerHisSerSerSerSerSerThrPheGlnPheSerSerSer 176

Qy        499 AGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATA 558
              |||   |||||||||:::|||||||||||||||||:::||||||:::   ||||||||
Db        177 SerSerGlnValGlySerGlySerAlaLeuProThrSerAspAsnAlaTyrArgGlyIle 196

Qy        559 CTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGC 618
              ||||||||||||||   ||||||||||||||   ||||||||
Db        197 LeuAsnProSerGlnLeuGlyGlnSerSerSerPheSerGlnThr--------------- 211

Qy        619 AGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCC 678
              ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 SerGlyGlnArgValSerSerAsnGlnArgProCysSerSerAspIleProAspSerPro 231

Qy        679 TGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTG 738
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        232 CysSerGlyGlyProIleIleSerHisSerGlyProTyrIleProSerSerHisSerVal 251

Qy        739 TCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGA 798
              ||||||||||||||||||||||||||||||:::|||||||||||||||   ||||||
Db        252 SerGlyGlyGlnArgProValValValValValGluGlnHisGlySerGlyGlyProGly 271

Qy        799 GTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATC 858
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        272 ValValGlnGlyLeuProCysSerAsnGlyGlyLeuProGlyLysProCysProProIle 291

Qy        859 ACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTAT 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSerAspSerTyr 311

Qy        919 CTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAA 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 LeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyrPheThrLys 331

Qy        979 GAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCT 1038
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 AspAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyProProIleSer 351

Qy       1039 GAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCC 1098
              ||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        352 GluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerGlyAlaSerSerVal 371

Qy       1099 ATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGC 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 IleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGlySerThrGly 391

Qy       1159 TCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGC 1218
              ||||||||||||||||||   ||||||||||   ||||||||||||||||||||
Db        392 SerLysGlyProCysSerProPheSerSerArgValHisSerSerSerSerIleSerSer 411

Qy       1219 AGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCA 1278
              |||:::|||||||||||||||||||||||   ||||||||||||||||||||
Db        412 SerSerGlySerProTyrHisProCysGlySerThrSerGlnSerProCysSerProPro 431

Qy       1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAG 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIleIleLeuGln 451

Qy       1339 CCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACA 1398
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 ProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSerSerLeuThr 471

Qy       1399 CTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGT 1458
                 |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 LeuThrGlyAspProAspGlySerProHisProAspProSerAlaGlyAlaLysProCys 491

Qy       1459 GGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTG 1518
                 ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GlySerGlySerAlaGlyLysIleProCysArgSerIleArgAspIleLeuAlaGlnVal 511

Qy       1519 AAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTC 1578
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 LysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGlyGluLeuLeu 531

Qy       1579 GACAGTCCA 1587
                 |||||||||
Db        532 AspSerPro 534

RESULT 4
CDSN_MOUSE
ID   CDSN_MOUSE     STANDARD;      PRT;   561 AA.
AC   Q7TPC1;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=Cdsn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C3H/He; TISSUE=Osteoblast;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BC055373; AAH55373.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039518; Mus musculus.
DR   MGI; MGI:3505689; Cdsn.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    561       Corneodesmosin.
FT   COMPBIAS      62    493       Ser-rich.
SQ   SEQUENCE   561 AA;  54414 MW;  7A6EBEF364038746 CRC64;

Alignment Scores:
Pred. No.:              3.05e-85        Length:         561
Score:                  1746.00         Matches:        362
Percent Similarity:     70.3%           Conservative:   49
Best Local Similarity:  61.9%           Mismatches:     90
Query Match:            56.3%           Indels:         84
DB:                     1               Gaps:           12

US-10-677-865A-2 (1-1590) x CDSN_MOUSE (1-561)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||:::||||||   ||||||   ||||||||||||||||:::|||||||||
Db          1 MetGlySerSerGlnAlaProArgMetGlySerValGlyGlyHisGlyLeuMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||:::|||||||||:::|||||||||   ||||||||||||||||||||   ||||||
Db         21 LeuMetAlaGlyLeuIleLeuProGlyIleLeuAlaLysSerIleGlyThrLeuSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuIleGlyLysThr 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
                 |||:::      ||||||      ||||||||||||
Db         61 GlySerAsnSerIleSerSerGlnGlyGlySerSerSerPhe----------------- 74

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC------GGATCCAGCATTGCC 294
                 ||||||   ||||||||||||      ||||||      |||||||||   :::
Db         75 ------SerSerGlnGlyGlySerSerSerPheSerSerHisGlyGlySerSerSerSer 92

Qy        295 CAGGGTGGTTCTGCAGGATCT------TTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
              ||||||   |||:::||||||      :::|||||||||||||||||||||||||   |||
Db         93 GlnGlySerSerSerGlySerLeuIleTyrLysProGlyThrGlyTyrSerGlnSerSer 112

Qy        349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGT----------------- 390
              ||||||   ||||||||   :::      |||   ||||||
Db        113 TyrSerTyrGlySerGlyGlySerArgProGlyGlySerGlySerGlnThrGlySerSer 132

Qy        391 TCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGC 450
              |||||||||   ||||||||||||||   ||||||||||||||||||   |||||||||
Db        133 SerSerGlnThrGlySerSerSerSerGlnSerGlySerSerGlySerGlnSerGlySer 152

Qy        451 AGCAGCTCTCATTCGAGCAGC---------------------------------------- 471
              |||   |||   |||   |||
Db        153 SerGlySerGlnSerGlySerSerGlySerGlnSerGlySerSerGlySerGlnSerGly 172

Qy        472 ----------------------------------------------------AGCAGCAGC 480
                                                                 |||||||||
Db        173 SerSerGlySerGlnSerGlySerSerGlySerGlnSerGlyArgTrpValSerSerSer 192

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAA----------------------------GTA 510
                 |||:::   |||||||||   |||
Db        193 SerGlnTrpValSerSerSerSerGlnSerGlySerSerGlySerSerArgAspArgPro 212

Qy        511 GGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCC 570
              |||:::||||||||||||||||   |||
Db        213 GlySerGlySerAlaLeuProThrGlyAsp---------------------------- 222

Qy        571 CAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGC 630
                     ::::::|||      ||||||   ||||||   ||||||:::   ||||||
Db        223 ---------LysThrSerGlyMetSerGln---SerGlyGlySerSerThrSerGlnSer 238

Qy        631 GTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGG 690
                 |||||||||   ||||||||||||||:::::|||||||||||||||||||||||||
Db        239 SerSerSerAsnLeuArgProCysSerSerAsnValProAspSerProCysSerGlyGly 258

Qy        691 CCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAG 750
              |||:::::::::|||||||||||||||||      :::|||:::|||||||||||||||
Db        259 ProValIleThrHisSerGlyProTyrIleSerGlyThrHisThrValSerGlyGlyGln 278
```

```
Qy        751 AGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGT 810
              ||||||||||||||||||||:::   |||||||||      ||||||      ||||||
Db        279 ArgProValValValValValGluHisHisGlySerAlaGlyProGly---PheGlnGly 297

Qy        811 CCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGAC 870
              ||||||||||||||||||      ||||||||||||||||||||||||||||||
Db        298 MetProCysSerAsnGlyGlyProAlaGlyLysProCysProProIleThrSerValGln 317

Qy        871 AAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGC 930
              |||   ||||||||||||||||||||||||||||||:::::|||||||||||||||||||
Db        318 LysProTyrGlyGlyTyrGluValValGlyGlySerAlaAsnSerTyrLeuValProGly 337

Qy        931 ATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTG 990
              ||||||||||||   ||||||||||||||||||||||||||||||:::||||||:::
Db        338 MetThrTyrSerGlyGlyLysIleTyrProValGlyTyrPheThrLysAspAsnProIle 357

Qy        991 AAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATAC 1050
              :::||||||||||||   ||||||||||||||||||||||:::||||||||||||||||
Db        358 ArgGlySerProGlyAlaProSerPheAlaAlaGlyProProValSerGluGlyLysTyr 377

Qy       1051 TTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCC---------- 1098
              ||||||||||||||||||||||||||:::   ::::::||||||:::
Db        378 PheSerSerAsnProIleIleProSerArgGlySerSerSerSerSerGlyTyrProVal 397

Qy       1099 ---ATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACG 1155
                 :::   ||||||||||||||:::||||||||||||   ||||||   ||||||
Db        398 GlyValValPheGlnProValGlySerGlyGlyValGlnProCysGlyThrGlySerVal 417

Qy       1156 GGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCC 1215
              ||||||||||||||||||   :::      ::::::   ||||||||||||   :::
Db        418 SerSerLysGlyProCysSerGlyThrArgIleGlnIleThrSerSerSerSerSerThr 437

Qy       1216 AGCAGCGCCGGTTCACCCTACCATCCCTGC---GGCAGTGCTTCCCAGAGCCCCTGCTCC 1272
              |||               ||||||||||||   |||   :::||||||   |||||||||
Db        438 Ser---------------TyrHisProCysSerGlyGlySerSerGlnGlyProCysSer 452

Qy       1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATC 1332
              ||||||||||||||||   |||      |||||||||   ||||||||||||||:::
Db        453 SerProGlyThrGlySerIleSerGlyGlySerSerSerLeuSerSerGlyLysIleVal 472

Qy       1333 CTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCC 1392
              |||||||||||||||||||||:::|||||||||:::||||||:::||||||   |||
Db        473 LeuGlnProCysGlySerLysSerThrSerSerGlyTyrProCysLeuSerValProSer 492

Qy       1393 TTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAG 1452
                 |||   ||||||   :::|||||||||   |||   ||||||   |||   |||
Db        493 SerProLeuAsnGlyGlyLeuAsnGlySerProGlnProValProSerValGlyValLys 512

Qy       1453 CCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCC 1512
              ||||||   :::|||   |||:::::|||||||||||||||||:::||||||
Db        513 LeuCysGlyLeuAsnSerProGlyArgValProCysArgSerIleArgAsnIleLeuThr 532

Qy       1513 CAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAG 1572
              ||||||||||||||||||||||||||   ||||||:::|||   ||||||||||||||||
Db        533 GlnValLysProLeuGlyProGlnLeuMetAspProLysValSerLeuProGlnGlyGlu 552

Qy       1573 TTACTCGACAGTCCA 1587
                 |||
Db        553 ProGlnGlyGluPro 557

RESULT 5
CDSN_PIG
ID   CDSN_PIG       STANDARD;      PRT;   225 AA.
AC   O19084;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin (S protein) (Fragment).
GN   Name=CDSN;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   TISSUE=Peripheral blood;
RA   Velten F.W., Rogel-Gaillard C., Renard C., Pontarotti P.,
RA   Tazi-Ahnini R., Vaiman M., Chardon P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; Z97402; CAB10690.1; -; Genomic_DNA.
FT   COMPBIAS     79     85       Poly-Ser.
FT   COMPBIAS    129    140       Poly-Ser.
FT   COMPBIAS    159    168       Poly-Ser.
FT   COMPBIAS    179    184       Poly-Ser.
FT   NON_TER       1      1
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  21644 MW;  8544070C41F7B280 CRC64;

Alignment Scores:
Pred. No.:                4.76e-44        Length:         225
Score:                    974.00          Matches:        186
Percent Similarity:       87.2%           Conservative:   11
Best Local Similarity:    82.3%           Mismatches:     19
Query Match:              31.4%           Indels:         10
DB:                       1               Gaps:           2

US-10-677-865A-2 (1-1590) x CDSN_PIG (1-225)

Qy        838 GGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTG 897
              ||||||||||||||||||||||||||||||||||||||||||   :::|||   ||||||||||
Db          1 GlyLysProCysProProIleThrSerValAspLys---PheGlySerTyrGluValVal 19

Qy        898 GGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db         20 GlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerGlyGlyLysIleTyr 39

Qy        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTT 1017
              ||||||||||||||||||||:::|||||||||||||||||||||||||   ||||||||
Db         40 ProValGlyTyrPheThrLysAspAsnProValLysGlySerProGlyAlaProSerPhe 59

Qy       1018 GCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 AlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSer 79

Qy       1078 --------------------------CAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAG 1110
                                       ||||||   ||||||||||||||   ||||||
Db         80 HisSerSerSerSerSerAsnIleTyrGlnSerGlyAlaSerSerAlaIleValPheGln 99

Qy       1111 CCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCC 1170
              |||||||||:::|||||||||||||   ||||||   ||||||:::|||||||||||||||
Db        100 ProValGlySerGlyGlyValGlnProCysGlyValGlySerSerGlySerLysGlyPro 119

Qy       1171 TGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCA 1230
              ||||||   |||   |||   |||   ||||||||||||||||||||:::||||||
Db        120 CysSerLeuSerGlySerGlyValHisSerSerSerSerIleSerSerSerSerGlySer 139

Qy       1231 CCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCC 1290
                 :::|||||||||||:::   ||||||   ||||||||||||||||||||||||||||
Db        140 SerPheHisProCysGlyThrValSerGlnGlyProCysSerProProGlyThrGlySer 159

Qy       1291 TTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
              ||||||   ||||||||||||:::|||   ||||||||||||||||||||||||||||||
Db        160 PheSerGlySerSerSerSerLysSerGlyGlyLysIleIleLeuGlnProCysGlySer 179

Qy       1351 AAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC 1410
```

```
                ||||||||||||||||||||||||||:::|||||||||||   ||||||:::||||||
Db        180 LysSerSerSerSerGlyHisProCysIleSerValSerSerSerThrLeuSerGlyGly 199

Qy       1411 CCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGT 1470
                |||||||||||||||   ||||||||||||||||||||||||||||||||||   |||
Db        200 ProAspGlySerProGlnProAspProSerAlaGlyAlaLysProCysGlySerGlySer 219

Qy       1471 GCTGGAAAGATCCCCTGC 1488
                :::|||||||||||||||
Db        220 SerGlyLysIleProCys 225

RESULT 6
Q767M8_PIG
ID   Q767M8_PIG  PRELIMINARY;      PRT;   262 AA.
AC   Q767M8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin.
GN   Name=CDSN;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=large white;
RA   Shigenari A., Ando A., Renard C., Chardon P., Shiina T., Kulski J.K.,
RA   Yasue H., Inoko H.;
RT   "Nucleotide sequencing analysis of the swine 433-kb genomic segment
RT   located between the non-classical and classical SLA class I gene
RT   clusters.";
RL   Immunogenetics 55:695-705(2004).
DR   EMBL; AB113354; BAD08421.1; -; Genomic_DNA.
SQ   SEQUENCE   262 AA;  25289 MW;  21F44CFE54C920F2 CRC64;

Alignment Scores:
Pred. No.:              3.43e-36        Length:         262
Score:                  827.00          Matches:        178
Percent Similarity:     60.8%           Conservative:   11
Best Local Similarity:  57.2%           Mismatches:     49
Query Match:            26.7%           Indels:         73
DB:                     2               Gaps:           3

US-10-677-865A-2 (1-1590) x Q767M8_PIG (1-262)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
                |||   |||   ||||||      :::|||||||   ||||||   ||||||:::||||||
Db          1 MetAspSerLeuArgAlaSerArgIleGlyArgAlaGlyGlyArgGlyMetLeuAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
                |||||||||   |||||||||   ||||||||||||||||||||||||||||||   |||
Db         21 LeuLeuAlaSerLeuLeuLeuGlnGlyThrLeuAlaLysSerIleGlyThrPheLeuAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
                |||||||||||||||||||||||   ||||||||||||||||||||   |||||||||
Db         41 ProCysLysAspProThrArgIleAlaSerProAsnAspProCysLeuLeuGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
                ||||||   |||||||||:::|||   ||||||||||||   ||||||||||||||||
Db         61 GlySerSerSerPheSerSerHisSerSerSerSerSerSerSerSerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
                      |||   |||||||||   |||   ||||||||||||              |||
Db         81 ProSerGlySerSerGlyGlySerHisGlyGlySerSerGlySerGlyGlyGlySerGly 100

Qy        301 GGTTCTGCAGGATCT---------------------------------------------- 315
                |||||||:::||||||
Db        101 GlySerSerGlySerSerSerGlySerSerGlyGlySerGlyGlySerIleAlaValHis 120

Qy        316 --------------------TTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354
                                   |||||||||||||||||||||||||||:::|||||||||

Db        121 SerGlySerSerGlySerLeuPheLysProGlyThrGlyTyrSerGlnIleSerTyrSer 140

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
                |||||||||||||||:::|||||||||||||||   ||||||||   ||||||   |||
Db        141 SerGlySerGlySerSerLeuGlnGlyAlaSerSerSerSerGlnSerGlySerIleSer 160

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
                |||   |||
Db        161 SerGlnSerAlaGly------------------------------------------- 165

Qy        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534

Db        165 ------------------------------------------------------------ 165

Qy        535 AATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594
                                                        |||   |||||||||
Db        166 ----------------------------------------GlyLeuSerSerSerVal 171

Qy        595 TCCCAAACCTCTGGGGTA---TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCC 651
                |||   ||||||   :::   |||||||||||||||   |||:::||||||   ||||||
Db        172 SerHisThrSerTrpMetSerSerSerSerGlyGlnLysValAsnSerAsnLeuArgPro 191

Qy        652 TGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGC 711
                ||||||||||:::|||||||||||||||||||||||||||||||||||   ||||||
Db        192 CysSerSerAspValProAspSerProCysSerGlyGlyProIleValSerGlnSerGly 211

Qy        712 CCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTG 771
                  ||||||   |||||||||||||||||||||||:::   ||| ||||||||||||||
Db        212 SerTyrIleSerSerSerHisSerValSerGlyGlyLysAsnPro-TrpTrpTrpTrpTr 231

Qy        772 GACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGC 831
                |:::||||||   ||||||      ||||||   |||   ||||||||||||      ||
Db        231 pSerSerThrAlaLeuValAlaProGluTrpAlaLysGlyProProValAlaAlaAlaAl 251

Qy        832 CTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
                |||||||||||||||||||||   ||||||
Db        251 aPheGlnAlaSerProValProProSerPro 261

RESULT 7
Q7Z559_HUMAN
ID   Q7Z559_HUMAN PRELIMINARY;      PRT;   133 AA.
AC   Q7Z559;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim Y.-K., Hur S.-S., Kim T.-Y., Kim T.-G., Pyo C.-W.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY328475; AAP94641.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  13126 MW;  5E6B2CA43FFDB73E CRC64;

Alignment Scores:
Pred. No.:              4.61e-29        Length:         133
Score:                  693.00          Matches:        131
Percent Similarity:     99.2%           Conservative:   1
Best Local Similarity:  98.5%           Mismatches:     1
Query Match:            22.3%           Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q7Z559_HUMAN (1-133)

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552
                |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
```

```
Db          1 SerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAspAsnSerTyrArg 20

Qy        553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGlyVal 40

Qy        613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         41 SerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIleProAsn 60

Qy        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSerHis 80

Qy        733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGlyAla 100

Qy        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCysPro 120

Qy        853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891
              |||||||||||||||||||||||||||||||||||||||
Db        121 ProIleThrSerValAspLysSerTyrGlyGlyTyrGlu 133

RESULT 8
Q7Z560_HUMAN
ID   Q7Z560_HUMAN PRELIMINARY;      PRT;   133 AA.
AC   Q7Z560;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim Y.-K., Hur S.-S., Kim T.-Y., Kim T.-G., Pyo C.-W.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY328474; AAP94640.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  13187 MW;  350E30DD5AFC652A CRC64;

Alignment Scores:
Pred. No.:              5.21e-29       Length:         133
Score:                  692.00         Matches:        131
Percent Similarity:     98.5%          Conservative:   0
Best Local Similarity:  98.5%          Mismatches:     2
Query Match:            22.3%          Indels:         0
DB:                     2              Gaps:           0

US-10-677-865A-2 (1-1590) x Q7Z560_HUMAN (1-133)

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db          1 SerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAspAsnSerTyrArg 20

Qy        553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         21 GlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrPheGlyVal 40

Qy        613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIleProAsp 60

Qy        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSerHis 80

Qy        733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGlyAla 100

Qy        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCysPro 120

Qy        853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891
              |||||||||||||||||||||||||||||||||||||||
Db        121 ProIleThrSerValAspLysSerTyrGlyGlyTyrGlu 133

RESULT 9
Q4FX61_LEIMA
ID   Q4FX61_LEIMA PRELIMINARY;      PRT;  2425 AA.
AC   Q4FX61;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Proteophosphoglycan ppg1.
GN   ORFNames=LMJ_0987;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
RX   PubMed=16020728; DOI=10.1126/science.1112680;
RA   Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA   Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA   Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
RA   Bianchettin G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA   Cadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
RA   Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA   Fazelina G., Fosker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
RA   Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
RA   Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA   Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA   Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA   Oliver K., O'Neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA   Quail M.A., Rabbinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA   Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA   Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA   Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA   Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA   Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA   Myler P.J.;
RT   "The Genome of the Kinetoplastid Parasite, Leishmania major.";
RL   Science 309:436-442(2005).
DR   EMBL; CP000081; AAZ14282.1; -; Genomic_DNA.
SQ   SEQUENCE   2425 AA;  225601 MW;  B3D24E1CB5262979 CRC64;

Alignment Scores:
Pred. No.:              7.27e-13       Length:         2425
Score:                  392.00         Matches:        170
Percent Similarity:     40.7%          Conservative:   52
Best Local Similarity:  31.2%          Mismatches:     268
Query Match:            12.6%          Indels:         55
DB:                     2              Gaps:           14

US-10-677-865A-2 (1-1590) x Q4FX61_LEIMA (1-2425)

Qy          9 GTCTCGGGCACCCTGGATGGGGCGTGTGGG--------------TGGGCACGGGATGAT 53
              ||||||      ::::::      ||||||                ||||||      :::
Db        571 ValSerLeuGluValAsnArgPheCysGlyCysLeuProSerThrTrpAlaAspLysAsn 590

Qy         54 GGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTT 113
              |||   ::: :::                |||      ||||||            |||
Db        591 GlyValSer-MetMetIleThrAlaGlnProGluLeuLeuAlaSerAspCysAlaThr-- 609

Qy        114 CTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCCTGCCTCACTGG 173
```

```
Db        610 -GluAsnAlaCysLysProGluThrGluSerSerSerSerAlaProSerAlaSerSerSe 629

Qy        174 GAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 233

Db        629 rSerAlaSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSe 649

Qy        234 TTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGC 293

Db        649 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAl 669

Qy        294 CCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTC 353

Db        669 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerSerAlaProSerAlaSe 689

Qy        354 CTCC------------GGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCT 401

Db        689 rSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaPr 709

Qy        402 GGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCA 461

Db        709 oSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSe 729

Qy        462 TTCG------------AGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGT 509

Db        729 rSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerSerAlaProSe 749

Qy        510 AGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTC 569

Db        749 rAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSe 769

Qy        570 CCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAG 629

Db        769 rAlaPro------SerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSe 787

Qy        630 CGTCAGCTCCAACCAGCGTCCC------TGTAGTTCGGACATCCCCGACTCTCCCTGCAG 683

Db        787 rSerSerSerSerSerAlaProSerAlaSerSerSerAlaProSerSerSerSerSe 807

Qy        684 TGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGG 743

Db        807 rSerAlaProSerAlaSerSerSer------AlaProSerSerSerSerSerAl 825

Qy        744 GGGTCAGAGGCCTGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTG-- 801

Db        825 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSe 845

Qy        802 ------------------------GTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCC 836

Db        845 rSerSerSerAlaSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaPr 865

Qy        837 A---GGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGT 893

Db        865 oSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSe 885

Qy        894 GGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAAT 953

Db        885 rSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAla----- 903

Qy        954 CTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTC 1013

Db        904 ------------------------------ProSerAlaSerSerSerSerAlaProSe 913

Qy       1014 CTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCC 1073

Db        913 rSerSerSerSerAlaProSerAlaSerSer---------SerSerAlaProSerSerSe 930

Qy       1074 CAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCA 1133

Db        930 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSer----- 948

Qy       1134 GCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGT 1193

Db        949 -------SerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAl 966

Qy       1194 CCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGC 1253

Db        966 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSe 986

Qy       1254 TTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCA 1313

Db        986 rSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSe 1006

Qy       1314 ATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCC 1373

Db       1006 rSerSerSerSer------SerAlaProSerAlaSerSerSerSerAlaProSerSerSe 1024

Qy       1374 TTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGA 1433

Db       1024 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAl 1044

Qy       1434 TCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTC 1493

Db       1044 aProSerAlaSerSerSerAlaProSerSerSerSerSerSer-AlaProSerAlaS 1064

Qy       1494 CATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGT 1553

Db       1064 erSerSer---Ser-AlaProSerSerSerSerSerAlaProSerAlaSerSerSerSer 1082

Qy       1554 TTTCCTACC 1562

Db       1083 AlaProSer 1085
```

RESULT 10
Q4FX64_LEIMA
ID   Q4FX64_LEIMA PRELIMINARY;   PRT; 1435 AA.
AC   Q4FX64;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Proteophosphoglycan ppg3, putative.
GN   ORFNames=LMJ_0984;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
RX   PubMed=16020728; DOI=10.1126/science.1112680;
RA   Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA   Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA   Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
RA   Bianchettin G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA   Cadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
RA   Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA   Fazelina G., Fosker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
RA   Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
RA   Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA   Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA   Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA   Oliver K., O'Neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA   Quail M.A., Rabbinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA   Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA   Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA   Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA   Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA   Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA   Myler P.J.;
RT   "The Genome of the Kinetoplastid Parasite, Leishmania major.";
RL   Science 309:436-442(2005).
DR   EMBL; CP000081; AAZ14279.1; -; Genomic_DNA.
SQ   SEQUENCE   1435 AA; 143303 MW; D0BCEEFF72B0C694 CRC64;

Alignment Scores:

Pred. No.:	1.2e-12	Length:	1435
Score:	387.50	Matches:	169

Percent Similarity: 42.1% Conservative: 65
Best Local Similarity: 30.4% Mismatches: 257
Query Match: 12.5% Indels: 65
DB: 2 Gaps: 15

US-10-677-865A-2 (1-1590) x Q4FX64_LEIMA (1-1435)

```
Qy          9 GTCTCGGGCACCCTGGATGGGGCGTGTGGG----------TGGGCACGGGATGAT 53
              ||||||      ::::::      ||||||          ||||||      :::
Db        618 ValSerLeuGluValAsnArgPheCysGlyCysLeuProSerThrTrpAlaAspLysAsn 637

Qy         54 GGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTT 113
              |||   ::: :::              |||      ||||||            |||
Db        638 GlyValSer-MetMetIleThrAlaGlnProGluLeuLeuAlaSerAspCysAlaThr-- 656

Qy        114 CTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGG 173
                  :::   ||||||      |||      :::|||                  :::
Db        657 -GluAsnAlaCysLysProGluThrGluSerSerSerAlaProSerAlaSerSerSe 676

Qy        174 GAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 233
                    ||||||      |||:::   |||   ||||||||||||      |||
Db        676 rSerAlaSerSerSerSerSerAlaProSerAlaSerSerSerAlaProSerSe 696

Qy        234 TTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT-- 291
                ||||||:::      |||          ||||||      ||||||   ||||||
Db        696 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaPr 716

Qy        292 ------------GCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTC 338
                        :::      |||:::   |||      |||              ||
Db        716 oSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSe 736

Qy        339 CCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCA 398
                |      ||||||   |||      :::          :::|||      |||
Db        736 rSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSe 756

Qy        399 GCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTC 458
                  |||:::   |||   |||   ||||||          |||   |||||||||::
Db        756 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAl 776

Qy        459 TCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGG 518
                :   |||:::|||||||||||||       |||||||||||          :::
Db        776 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSe 796

Qy        519 CTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGG 578
                 |||:::   |||::::::   :::|||                  |||   |||
Db        796 rSerSerAlaProSerSerSerSerSer--------------------SerAlaProSe 809

Qy        579 ACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTC 638
                  |||||||||||:::   :::|||      |||||||||      |||   |||||
Db        809 rAlaSerSerSerSerAlaProSerSer------SerSerSerAlaProSerAlaSerSe 827

Qy        639 CAACCAGCGTCCC------TGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCC 692
                |:::      |||         ||||||      |||   :::      |||      ||
Db        827 rSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaPr 847

Qy        693 CATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC-------------------- 732
                |      |||   |||   |||          ||||||
Db        847 oSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSe 867

Qy        733 -------TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTC 785
                      |||   |||          |||                      ||
Db        867 rAlaProSerAlaSerSerSerSerAlaPro--------------SerSerSerSerSe 882

Qy        786 TGGTGCCCCTGGAGTG------GTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCA-- 837
                |   ||||||            |||      |||:::         |||
Db        882 rSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSe 902

Qy        838 ----GGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGT 893
                        |||   :::|||          :::
Db        902 rAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSe 922

Qy        894 GGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAAT 953
                  ||||||   |||        |||      :::   |||
```

```
Db        922 rSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaPro-- 941

Qy        954 CTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTC 1013
                              :::         |||         |||         |||||
Db        942 ------------SerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSe 957

Qy       1014 CTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCC- 1072
                |   :::::       ||||||      ::::::::::|||::::::   ||||||
Db        957 rSerSerSer-ProAlaProSerAlaSerSerSerAlaProSerSerSerSerSerA 977

Qy       1073 --CCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGT 1130
                  |||::::::          |||      ||||||          ||||||  |
Db        977 laProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProLeuValSerS 997

Qy       1131 CCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCG 1190
                |||||||      :::  |||:::   |||      |||   :::   |||||||||
Db        997 erSerSerAlaProSer--SerSerSerSerSerAlaProSerAlaSerSerSerSerAl 1016

Qy       1191 AGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAG 1250
                      ||||||||||||   :::   ||||||   |||          |||      ||
Db       1016 aProSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSe 1036

Qy       1251 TGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTC 1310
                |:::|||   :::|||   :::          :::          ||||||||||||::
Db       1036 rSerSerSerAlaProSerAlaSerSerSerAlaProSerSerSerSerSerSerAl 1056

Qy       1311 CCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCA 1370
                :   |||:::                          :::   ||||||||||||
Db       1056 aProSerAlaSerSer--------------SerSerAlaProSerSerSerSerSerAl 1071

Qy       1371 CCCTTGCATGTCTGTCTCCTCCTTG------ACACTGACTGGGGGCCCCGATGGCTCTCC 1424
                |||      |||   ||||||          :::   :::      |||      |||
Db       1071 aProSerAlaSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSe 1091

Qy       1425 CCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCC 1484
                          ||||||:::   :::          |||:::|||:::      :::|
Db       1091 rSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSer-AlaP 1111

Qy       1485 CTGCCGCTCCATCCGGGATATCCTAGC-----------------CCAAGTGAAGCCTCT 1526
                ||::::::   |||       ||||||                 ||||||   :::|||
Db       1111 roSerSerSerSerSerAla-ProSerAlaSerSerSerSerAlaProSerSerSerSer 1130

Qy       1527 GGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACC 1562
                  ||||||::::::          |||   |||:::
Db       1131 SerAlaProSerAlaSerSerSerSerAlaProSer 1142
```

RESULT 11
Q4FX62_LEIMA
ID Q4FX62_LEIMA PRELIMINARY; PRT; 17392 AA.
AC Q4FX62;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Proteophosphoglycan 5.
GN ORFNames=LMJ_0986;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RX PubMed=16020728; DOI=10.1126/science.1112680;
RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
RA Bianchettin G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA Cadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
RA Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA Fazelina G., Fosker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
RA Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,

```
RA   Oliver K., O'Neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA   Quail M.A., Rabbinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA   Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA   Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA   Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA   Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA   Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA   Myler P.J.;
RT   "The Genome of the Kinetoplastid Parasite, Leishmania major.";
RL   Science 309:436-442(2005).
DR   EMBL; CP000081; AAZ14281.1; -; Genomic_DNA.
SQ   SEQUENCE   17392 AA;  1505876 MW;  E693BFC2CD9CAF70 CRC64;

Alignment Scores:
Pred. No.:                1.53e-12      Length:          17392
Score:                    387.50        Matches:         167
Percent Similarity:       42.4%         Conservative:    55
Best Local Similarity:    31.9%         Mismatches:      262
Query Match:              12.5%         Indels:          41
DB:                       2             Gaps:            10

US-10-677-865A-2 (1-1590) x Q4FX62_LEIMA (1-17392)

Qy         68 CTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTA 127
Db       8176 LeuAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSer 8195

Qy        128 AGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACT-GGGAAGGGTGACTCC 186
Db       8196 SerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSer 8215

Qy        187 AGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGA 246
Db       8216 SerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAla 8235

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCT 306
Db       8236 ProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSer 8255

Qy        307 GCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGC 366
Db       8256 SerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAla 8275

Qy        367 TCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGA 426
Db       8276 ProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSer 8295

Qy        427 AGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAG 486
Db       8296 SerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaPro 8315

Qy        487 TTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCT 546
Db       8316 SerAlaSerSerSerSerAlaProSerSerSerSerAlaProSerAlaSerSerSer 8335

Qy        547 TAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCT 606
Db       8336 SerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSer 8355

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCC------TGTAGTTCG 660
Db       8356 ------SerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSer 8373

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db       8374 SerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAla 8393

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db       8394 SerSerSerSerAlaProSerSerSerSerSer------------AlaProSerAlaSer 8409

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAA------GGTCCCCCCTGTAGCAATGGTGGCCTT 834
                                                                        ^
Db       8410 SerSerSerAlaProSerSerSerSerSerThrAlaProSerAlaSerSerSerSerAla 8429

Qy        835 CCA------GGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
Db       8430 ProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSer 8449

Qy        889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTG---GTTCCAGGCATGACCTACAGTAAG 945
Db       8450 SerAlaProSerAlaSerSerSerSerAlaLeuSerSerSerSerSerThrAlaProSer 8469

Qy        946 GGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGG 1005
Db       8470 GlySerSerSerSerAlaProSerSerSerSerSerAlaProSerGlySerSerSerSer 8489

Qy       1006 GTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCC 1065
Db       8490 AlaProSerSerSerSer-SerAlaProSerAlaSerSerSerSerAlaProSerSerSe 8509

Qy       1066 ATCATCC------------------------------CCAGCCAGTCGGCAGCTTCCTCG 1095
Db       8509 rSerSerThrAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAl 8529

Qy       1096 GCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACG 1155
Db       8529 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAla--Ser 8548

Qy       1156 GGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCC 1215
Db       8549 SerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAla 8568

Qy       1216 AGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCC--- 1272
Db       8569 ProSerSerSerSerSerThrAlaProSerGlySerSerSerSerAlaProSerSerSer 8588

Qy       1273 ------------CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       8589 SerSerSerAlaProSerGlySerSerSerSerAlaProSerSerSerSerSerSerAla 8608

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       8609 ------------------ProSerAlaSerSerSerSerAlaProSerSerSerSerSer 8622

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       8623 SerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSer 8642

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       8643 AlaSerSerSerSerAlaProSerSerSerSerSerSer-AlaProSerAlaSerSerSe 8662

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db       8662 r---Ser-AlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProS 8681

Qy       1561 CC 1562
Db       8681 er 8681

RESULT 12
Q4FX63_LEIMA
ID   Q4FX63_LEIMA PRELIMINARY;      PRT;  7194 AA.
AC   Q4FX63;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Proteophosphoglycan ppg4.
GN   ORFNames=LMJ_0985;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
```

```
RX   PubMed=16020728; DOI=10.1126/science.1112680;
RA   Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA   Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA   Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
RA   Bianchettin G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA   Cadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
RA   Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA   Fazelina G., Fosker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
RA   Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
RA   Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA   Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA   Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA   Oliver K., O'Neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA   Quail M.A., Rabbinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA   Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA   Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA   Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA   Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA   Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA   Myler P.J.;
RT   "The Genome of the Kinetoplastid Parasite, Leishmania major.";
RL   Science 309:436-442(2005).
DR   EMBL; CP000081; AAZ14280.1; -; Genomic_DNA.
SQ   SEQUENCE   7194 AA;  626431 MW;  9BFCBE646EAB554A CRC64;

Alignment Scores:
Pred. No.:              2.76e-12        Length:         7194
Score:                  382.00          Matches:        168
Percent Similarity:     40.5%           Conservative:   48
Best Local Similarity:  31.5%           Mismatches:     235
Query Match:            12.3%           Indels:         82
DB:                     2               Gaps:           16

US-10-677-865A-2 (1-1590) x Q4FX63_LEIMA (1-7194)

Qy         74 TCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACC 133
              ||||||         |||         :::   |||||||||   :::
Db       3328 SerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAla 3347

Qy        134 CCACGCGTATCACCTCCCCTA-----------------ACGACCCCTGCCTCACTGGGA 175
              |||      |||                       :::   |||:::|||
Db       3348 ProSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSer 3367

Qy        176 AGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTT 235
                  |||   |||  ||||||   |||      |||||||||   ||||||
Db       3368 SerSerAlaProLeuAla--SerSerSerSerAlaProSerSerSerSerSerSerAlaP 3387

Qy        236 CCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCC---------AGCGGATCCA 286
              ||||||  ||||||         ||||||  ||||||         |||  ||||
Db       3387 roSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerS 3407

Qy        287 GCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCA 346
              ||   :::        |||:::  |||      |||           |||
Db       3407 erSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaP 3427

Qy        347 GCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGA 406
                 ||||||  |||  :::  |||      ||||||  ||||||         |
Db       3427 roSerSerSerSerSerSerAlaProLeu------AlaSerSerSerSerAlaProSerS 3445

Qy        407 GCAGCAGC---------TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCT 457
              ||||||||         |||  |||  ||||||         |||  |||||||||:
Db       3445 erSerSerThrAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerA 3465

Qy        458 CTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATG 517
              ::   |||:::|||||||||||||       ||||||||||||          :::
Db       3465 laProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerS 3485

Qy        518 GCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTG 577
                ||:::   |||::::::   :::|||                |||   |||
Db       3485 erSerSerAlaProSerSerSerSerSer--------------------SerAlaProS 3498

Qy        578 GACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCT 637
                    ||||||||||||:::   :::|||      |||||||||      |||   ||||

Db       3498 erAlaSerSerSerSerAlaProSerSerSerSerSerSerSerAlaProSerAlaSerS 3518

Qy        638 CCAACCAGCGTCCC------TGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGC 691
              ||:::      |||         ||||||      |||   :::      |||      |
Db       3518 erSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaP 3538

Qy        692 CCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGA 751
              ||      |||  |||  |||         ||||||   :::  |||
Db       3538 roSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSer------- 3555

Qy        752 GGCCTGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTG---GTTCAAG 808
                                          |||  ||||||
Db       3556 ----------------------------SerSerSerAlaProSerAlaSerSerSerS 3566

Qy        809 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 868
                |||  |||:::        |||                                |||
Db       3566 erAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerS 3586

Qy        869 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 928
                |||                                  |||:::  |||      |||
Db       3586 erSerSer--------------------AlaProSerAlaSerSerSerAlaProS 3599

Qy        929 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 988
                :::  |||                                             |||
Db       3599 erSerSerSerSerAla-------------------------------------ProS 3606

Qy        989 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 1048
                  |||         ||||||   ::::::      |||   :::
Db       3606 erAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSer---- 3624

Qy       1049 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 1108
                |||||   |||         |||   ||||||   |||::::::   :::
Db       3625 -----SerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerA 3643

Qy       1109 AGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGAC 1168
                |||      :::                                 |||      |||
Db       3643 laProSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerS 3663

Qy       1169 CCTGCTCTCCC------TCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
                :::|||      |||||||||      |||||||||||||||   :::   ||||
Db       3663 erSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerA 3683

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCC---------------T 1267
                ||  |||   :::   |||      |||:::|||   :::|||
Db       3683 laSerSerSerPheAlaProSerSerSerSerSerAlaProSerAlaSerSerSerS 3703

Qy       1268 GCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAA 1327
                :::|||      :::   |||      |||:::|||||||||   :::
Db       3703 erAlaProSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerS 3723

Qy       1328 TCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCT 1387
                     |||      |||   ||||||:::                     |||
Db       3723 erSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaP 3743

Qy       1388 CCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT---CCTGATCCCTCCGCTG 1444
                |||   :::   :::      |||      |||      |||      |||:::
Db       3743 roSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAlaProSerAlaSerS 3763

Qy       1445 GTGCCAAGCCCTGTGGCTCCAGCAGTGCT-------------------------GGAA 1477
                :::   |||      ||||||||||||
Db       3763 erSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerS 3783

Qy       1478 AGATCCCCTGCCGCTCCATCCGGGATATCCTAGC------CCAAGTGAAGCCTCTGGGGC 1531
                |||   :::|||||||||      ||| |||      ||||||   :::|||   ||
Db       3783 er-SerSerSerAlaProSerAlaSerSer-SerSerAlaProSerSerSerSerSerAl 3802

Qy       1532 CCCAGCTAGCTGACCCTGAAGTTTTCCTACC 1562
                ||||::::::      |||   |||:::
Db       3802 aProSerAlaSerSerSerSerAlaProSer 3812

RESULT 13
Q9BKV7_LEIMA
ID   Q9BKV7_LEIMA PRELIMINARY;      PRT;  1325 AA.
```

```
AC   Q9BKV7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ppg3.
GN   Name=ppg3;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
RA   Myler P.J., Sisk E., Cawthra J., Handley F., Vogt C., Robertson L.,
RA   McDonagh P., Ivens A., Nguyen D., Munden H., Stuart K., Worthey E.A.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC084329; AAK31375.1; -; Genomic_DNA.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR007090; LRR_plant.
DR   Pfam; PF00560; LRR_1; 8.
DR   PRINTS; PR00019; LEURICHRPT.
KW   Leucine-rich repeat; Repeat.
SQ   SEQUENCE   1325 AA;  133856 MW;  8898928BAEE9418F CRC64;

Alignment Scores:
Pred. No.:              5.22e-12        Length:         1325
Score:                  375.50          Matches:        170
Percent Similarity:     41.1%           Conservative:   51
Best Local Similarity:  31.6%           Mismatches:     237
Query Match:            12.1%           Indels:         80
DB:                     2               Gaps:           19

US-10-677-865A-2 (1-1590) x Q9BKV7_LEIMA (1-1325)

Qy          9 GTCTCGGGCACCCTGGATGGGGCGTGTGGG--------------TGGGCACGGGATGAT 53
              ||||||      ::::::    ||||||                ||||||     :::
Db        618 ValSerLeuGluValAsnArgPheCysGlyCysLeuProSerThrTrpAlaAspLysAsn 637

Qy         54 GGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTT 113
              |||   ::: :::               |||      ||||||             |||
Db        638 GlyValSer-MetMetIleThrAlaGlnProGluLeuLeuAlaSerAspCysAlaThr-- 656

Qy        114 CTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGG 173
                 :::   ||||||     |||    :::|||                    :::
Db        657 -GluAsnAlaCysLysProGluThrGluSerSerSerSerAlaProSerAlaSerSerSe 676

Qy        174 GAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 233
                  ||||||      |||:::   |||   ||||||||||||      |||
Db        676 rSerAlaSerSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSe 696

Qy        234 TTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT-- 291
               ||||||:::      |||          ||||||      ||||||   ||||||
Db        696 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaPr 716

Qy        292 ------------GCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTC 338
                         :::         |||:::   |||      |||             ||
Db        716 oSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSe 736

Qy        339 CCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCA 398
              |              ||||||   |||      :::         :::|||      |||
Db        736 rSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSe 756

Qy        399 GCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTC 458
                 |||:::   |||   |||   ||||||          |||   |||||||||::
Db        756 rSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerSerAl 776

Qy        459 TCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGG 518
               :   |||:::|||||||||||||         ||||||||||||              :::
Db        776 aProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSe 796

Qy        519 CTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGG 578
                 |||:::   |||::::::   :::|||                  |||   |||
Db        796 rSerSerAlaProSerSerSerSerSer--------------------SerAlaProSe 809

Qy        579 ACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTC 638
                  ||||||||||||:::   :::|||      |||||||||      |||   |||||
Db        809 rAlaSerSerSerSerAlaProSerSer------SerSerSerAlaProSerAlaSerSe 827

Qy        639 CAACCAGCGTCCC---TGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCAT 695
              |:::      |||      ||||||      |||   :::      |||      |||
Db        827 rSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSe 847

Qy        696 CGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCC 755
                  |||   |||   |||          ||||||   :::   |||
Db        847 rSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSer----------- 863

Qy        756 TGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTG---GTTCAAGGTCC 812
                                              |||   ||||||
Db        864 ------------------------SerSerSerAlaProSerAlaSerSerSerSerAl 875

Qy        813 CCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAA 872
               |||   |||:::          |||                          |||
Db        875 aProSerSerSerSerSerAlaPro------------------------SerAlaSerSe 887

Qy        873 ATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCAT 932
               |||                          ||||||   |||          |||
Db        887 rSer------------------SerAlaProSerSerSerSerSerAlaProSerAl 901

Qy        933 GACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAA 992
                :::   |||                                          |||
Db        901 aSerSerSerSerAla-----------------------------------ProSerSe 909

Qy        993 AGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTT 1052
                  |||         ||||||   ::::::            |||
Db        909 rSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSer----- 927

Qy       1053 CTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCC 1112
               ||||||   |||         |||   ||||||   ||||||:::   :::      ||
Db        928 -SerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaPr 947

Qy       1113 AGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTG 1172
               |      :::                          |||:::   |||      |||
Db        947 oSerAlaSerSerSer------------SerAlaProSerSerSerSerSerAlaPro-- 962

Qy       1173 CTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACC 1232
               |||   |||||||||       ||||||||||||||   :::   ||||||   |||
Db        963 -SerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAlaSerSerSe 982

Qy       1233 CTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTT 1292
                  |||       |||:::|||   :::|||   :::       :::   |||
Db        982 rSerAlaProSerSerSerSerSerAlaProSerAla------SerSerSerAl 1000

Qy       1293 CAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAA 1352
               |||||||||||||||      |||                  |||      |||
Db       1000 aProSerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSe 1020

Qy       1353 GTCC------------AGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC------TCCTT 1394
                  |||            |||||||||       |||      |||   |||      |||
Db       1020 rSerAlaProSerAlaSerSerSerSerAlaProSerSerSerSerSerAlaProSerAl 1040

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
                :::   :::      |||      |||             |||||||||   :::
Db       1040 aSerSerSerSerAlaProSerSerSerSerSerAla---ProSerAlaSerSerSerSe 1059

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
                  |||||||||:::       :::|||:::|||   |||      ||| :::|||
Db       1059 rAlaProSerSerSerSerSer-AlaProSerAlaSerSerSer---Ser-AlaPro 1077

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACC 1562
                  |||   :::|||   ||||||::::::          |||   |||:::
Db       1078 SerSerSerSerSerAlaProSerAlaSerSerSerSerAlaProSer 1093

RESULT 14
Q7UE67_RHOBA
ID   Q7UE67_RHOBA PRELIMINARY;      PRT;   669 AA.
AC   Q7UE67;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
```

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RB11540;
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
OX   NCBI_TaxID=117;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294153; CAD79181.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   669 AA;  63517 MW;  F71218F7A659FC6C CRC64;

Alignment Scores:
Pred. No.:              9.94e-11      Length:          669
Score:                  351.00        Matches:         149
Percent Similarity:     41.1%         Conservative:    31
Best Local Similarity:  34.0%         Mismatches:      158
Query Match:            11.3%         Indels:          100
DB:                     2             Gaps:            13

US-10-677-865A-2 (1-1590) x Q7UE67_RHOBA (1-669)

Qy      178 GGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCC 237
Db       14 GlySerSerSerSerSerSerSerAspSerSerSerSerSerSerSerSerSerProSer 33

Qy      238 AGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAG 297
Db       34 SerSerSerSerSerGlySerSerSerSerSerSerSerGlySerSerSerSerSerSerMet 53

Qy      298 GGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCC 357
Db       54 SerGlySerSerSerSerSerMetSerSerSerSerSerSerLeuSerSerSerSer 73

Qy      358 GGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCT 417
Db       74 SerSerSerSer------------GlySerSerSerSerSerSerGlySerSerSerSer 89

Qy      418 CACTCGGGA------AGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCG------ 465
Db       90 SerSerGlySerProSerSerSerSerSerSerSerGlySerSerSerSerSerSerSerGly 109

Qy      466 AGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCT 525
Db      110 SerSerSerSerSerSerLysSerSerSerSerSerAspSerSerSerAsnSerSerSer 129

Qy      526 CTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGC 585
Db      130 SerAspSerSerAspSerSer--------------HisSerSerSerAlaSerSerSer 144

Qy      586 TCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAG 645
Db      145 SerAspSerSerAspSerSer------ThrSerSerSerSerSerValSerSerSerGln 162

Qy      646 CGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGG---CCCATCGTCTCG 702
Db      163 SerSerAspSerSerAspSerSerAspSerSerThrSerThrSerLeuProSerValSer 182

Qy      703 CACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTG 762
Db      183 SerSerSerSerSerAspGlySerSerSerSerSerSerGly----------------- 196

Qy      763 GTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGC 822

                                  |||   :::           |||           |||
Db      197 ------------------SerSerAspSerSerSerAsnSerGlnSerSerAspSerSer 210

Qy      823 AATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGT 882
            |||                                                 |||   |||
Db      211 AsnSer-----------------------------------------AspSerSer------ 215

Qy      883 GGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGT 942
                                  ||||||||||||      :::|||   :::
Db      216 ------------------AspSerSerAspSerThrSerLeuProSerIle--------- 226

Qy      943 AAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCA 1002
Db      226 ------------------------------------------------------------ 226

Qy     1003 GGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAAC 1062
                                  :::|||         |||:::|||         ||||||:::
Db      227 ------------------SerGlySerSerSerSerAspGlySerSerAlaSerSerAsp 240

Qy     1063 CCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACT 1122
                        |||   ::::::   ||||||
Db      241 SerSerAspSerSerSerAsnSerGlnSerSer--------------------------- 251

Qy     1123 GGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCC 1182
                                        |||:::   |||:::         |||
Db      252 ------------------------AspSerSerAsnSerGluSerSerAspSerSerAsp 263

Qy     1183 AGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCC 1242
            |||      :::||||||   |||      ||||||||||||||||||
Db      264 SerIleSerLeuProSerIleSerGlySerSerSerSerAlaGlySer---------Ser 280

Qy     1243 TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGC 1302
                  |||:::      |||      |||                 |||   ||||||:::
Db      281 SerAlaSerSerAspSerSerAsnSerSerSerAsnSerGlnSerSerAspSerSerThr 300

Qy     1303 TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCT 1362
            |||   |||      |||          :::      |||      |||   |||||||||
Db      301 SerGluSerAsnAspSerSerAlaSerValSerLeuProSerValSerGlySerSerSer 320

Qy     1363 TCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC------CCCGAT 1416
            |||                |||      |||   ::         |||          |||
Db      321 SerAlaSerAspSerSerSerSerGluSerGlnSerSerAspSerGlySerSerSerAsp 340

Qy     1417 GGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGT 1470
               |||                :::   :::   |||      |||||||||
Db      341 SerSerThrSerGlnSerSerGluSerAlaSerThrProSerValSerSerSer 358

RESULT 15
Q75JC9_DICDI
ID   Q75JC9_DICDI PRELIMINARY;      PRT;   374 AA.
AC   Q75JC9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Similar to Leishmania major. Ppg3 (Hypothetical protein).
GN   ORFNames=DDB0168484;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RX   MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA   Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA   Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA   Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL   Nature 418:79-85(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Baumgart C.;
```

```
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
DR   EMBL; AC117072; AAS38911.1; -; Genomic_DNA.
DR   EMBL; AAFI01000015; EAL71525.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   374 AA;  34423 MW;  227F4652789816DC CRC64;

Alignment Scores:
Pred. No.:              3.21e-10        Length:         374
Score:                  341.00          Matches:        141
Percent Similarity:     36.5%           Conservative:   35
Best Local Similarity:  29.3%           Mismatches:     170
Query Match:            11.0%           Indels:         136
DB:                     2               Gaps:           7

US-10-677-865A-2 (1-1590) x Q75JC9_DICDI (1-374)

Qy         73 CTCCTCCTGCCAGGGACCTTGGCT---------AAGAGCATTGGC---ACCTTCTCAGAC 120
              |||:::||||||      :::            ||||||:::      |||      :::
Db         15 LeuValLeuProSerIleValThrProTyrArgLysSerValGluIleThrAsnGluGlu 34

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAG--- 177
              |||   :::|||     ::::::         |||              |||
Db         35 ProGlnArgAspIleTyrGlnLeuGluLysThrAsnLysMetAlaGluValGlyAspLeu 54

Qy        178 ----------------------GGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCC 213
                                     |||:::|||      ||||||   |||   |||
Db         55 GlyValAlaSerPheLeuAsnIleLeuAspThrSerSerSerSerSerSerSerSerSer 74

Qy        214 AGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGC 273
              |||||||||   ||||||   ||||||:::   ||||||         ||||||   |||
Db         75 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 94

Qy        274 TCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGG 333
              ||||||   ||||||  :::        |||:::   |||          :::
Db         95 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 114

Qy        334 TATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCC 393
                 |||      |||   ||||||   |||   |||:::         :::|||   |||
Db        115 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 134

Qy        394 TCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGC 453
              |||          ||||||||||||   |||   ||||||   |||   |||   ||||||
Db        135 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 154

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGG 513
              ||||||   ||||||||||||||||||         ||||||||||||
Db        155 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 174

Qy        514 AATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAG 573
              :::   |||:::      ::::::   :::|||                      |||

Db        175 SerSerSerSerSerSerSerSerSerSerSer----------------------SerSer 187

Qy        574 CCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTC 633
                       ||||||||||||||||   :::|||      ||||||||      |||
Db        188 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 207

Qy        634 AGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCC 693
              ||||||:::           ||||||             |||      |||
Db        208 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 227

Qy        694 ATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGG 753
                  |||   |||                ||||||   |||   |||
Db        228 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer------------ 243

Qy        754 CCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCC 813

Db        243 ------------------------------------------------------------ 243

Qy        814 CCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAA 873

Db        243 ------------------------------------------------------------ 243

Qy        874 TCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATG 933
                                           ||||||   |||
Db        244 ----------------------------SerSerSerSer------------------ 247

Qy        934 ACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAA 993

Db        247 ------------------------------------------------------------ 247

Qy        994 GGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTC 1053

Db        247 ------------------------------------------------------------ 247

Qy       1054 TCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCA 1113
              ||||||:::           |||   |||::::::||||||:::
Db        248 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSer-------------- 262

Qy       1114 GTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGC 1173
                                                  |||:::   |||
Db        263 ------------------------------SerSerSerSerSerSerSerSer 271

Qy       1174 TCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCC 1233
              |||   |||||||||          ||||||||||||   |||||||||:::   |||
Db        272 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 291

Qy       1234 TACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTC 1293
                        |||:::|||   |||      |||          :::   |||
Db        292 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 311

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAG 1353
              ||||||||||||||||   ||||||                              |||
Db        312 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 331

Qy       1354 TCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCC 1413
              ||||||||||||               |||   ||||||   :::   :::
Db        332 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 351

Qy       1414 GATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
                   |||                ||:::   :::            |||||||||:::
Db        352 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 371

Qy       1474 GGAAAG 1479
                 |||:::
Db        372 GlyGlu 373


Search completed: March 21, 2006, 22:06:29
Job time : 271 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 21:53:09 ; Search time 29.4 Seconds
(without alignments)
4752.464 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152016_26056/app_query.fasta_1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_157_@runat_21032006_152016_26056 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2738	88.3	529	2	AAW80498	Aaw80498 A protein
2	2738	88.3	529	2	AAW85055	Aaw85055 Human pro
3	2738	88.3	529	4	AAU27550	Aau27550 Corneodes
4	2738	88.3	529	7	ABJ39148	Abj39148 Human cor
5	2738	88.3	529	8	ADN04226	Adn04226 Antipsori
6	2735	88.2	529	8	ADP12525	Adp12525 Protein e
7	2720	87.7	529	3	AAY44989	Aay44989 Human epi
8	2683	86.5	529	4	AAG64152	Aag64152 Polypepti
9	2491	80.3	486	7	ABJ39147	Abj39147 Human S p
10	2491	80.3	486	8	ADQ21504	Adq21504 Human sof
11	1092	35.2	219	8	ABO60412	Abo60412 Human gen
12	703	22.7	146	7	ABJ39195	Abj39195 Fragment
13	530	17.1	111	7	ABJ39196	Abj39196 Fragment
14	400	12.9	76	7	ABJ39197	Abj39197 Fragment
15	369	11.9	71	5	AAU91255	Aau91255 Nucleic a
16	335.5	10.8	950	4	ABB71513	Abb71513 Drosophil
17	332.5	10.7	914	9	ADV42115	Adv42115 Fruitfly
18	332.5	10.7	914	9	ADV44212	Adv44212 Drosophil
19	315	10.2	2112	4	ABB60403	Abb60403 Drosophil
20	310	10.0	58	7	ABJ39194	Abj39194 Fragment
21	299.5	9.7	763	2	AAW31852	Aaw31852 Mycobacte
22	298	9.6	55	7	ABJ39193	Abj39193 Fragment
23	294.5	9.5	572	2	AAW31855	Aaw31855 Mycobacte
24	290	9.3	2280	4	ABB61650	Abb61650 Drosophil
25	285	9.2	697	8	ADX72179	Adx72179 Plant ful
26	284.5	9.2	1063	5	ABB54168	Abb54168 Lactococc
27	265.5	8.6	842	4	ABB66631	Abb66631 Drosophil
28	265.5	8.6	864	4	ABB71319	Abb71319 Drosophil
29	263	8.5	258	7	ADB64880	Adb64880 Human pro
30	262.5	8.5	1010	2	AAW34989	Aaw34989 Teredinib
31	262	8.4	400	6	AAO16495	Aao16495 Kukulcani
32	260	8.4	693	5	ABP69529	Abp69529 Human pol
33	259.5	8.4	866	8	ADR86404	Adr86404 Aspergill
34	259	8.3	1459	8	ADN23765	Adn23765 Bacterial
35	257	8.3	559	8	ADQ67157	Adq67157 Novel hum
36	256.5	8.3	693	7	ADM05164	Adm05164 Human pro
37	256	8.3	386	9	ADZ47759	Adz47759 Mycobacte
38	253.5	8.2	1127	4	AAB95541	Aab95541 Human pro
39	253.5	8.2	1480	7	ADJ70433	Adj70433 Human hea
40	253.5	8.2	1828	4	ABB50164	Abb50164 Human tra
41	253.5	8.2	2427	8	ADR09758	Adr09758 Human pro
42	253.5	8.2	2735	6	ABR41356	Abr41356 Human DIT
43	253.5	8.2	2971	3	AAB41231	Aab41231 Human ORF
44	253	8.2	598	6	ABU37030	Abu37030 Protein e
45	252	8.1	2586	4	ABB66878	Abb66878 Drosophil

ALIGNMENTS

RESULT 1
AAW80498
ID AAW80498 standard; protein; 529 AA.
XX
AC AAW80498;
XX
DT 10-FEB-1999 (first entry)
XX
DE A protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761362-A1.
XX
PD 02-OCT-1998.
XX
PF 28-MAR-1997; 97FR-00003899.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586093/50.
DR N-PSDB; AAV63614.
XX
PT New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic

```
PT   or therapeutic treatment of trophic disorders or thinning of the
PT   epidermis.
XX
PS   Claim 1; Page 29; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              5.75e-172       Length:         529
Score:                  2738.00         Matches:        527
Percent Similarity:     99.8%           Conservative:   1
Best Local Similarity:  99.6%           Mismatches:     1
Query Match:            88.3%           Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x AAW80498 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529

RESULT 2
AAW85055
ID   AAW85055 standard; protein; 529 AA.
XX
AC   AAW85055;
XX
DT   05-FEB-1999  (first entry)
XX
DE   Human protein designated corneodesmosin.
XX
```

KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761363-A1.
XX
PD 02-OCT-1998.
XX
PF 11-SEP-1997; 97FR-00011317.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586094/50.
DR N-PSDB; AAV63541.
XX
PT New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT epidermis etc., also compositions containing protease that degrades
PT corneodesmosin.
XX
PS Claim 1; Page 30; 35pp; French.
XX
CC The present sequence represents a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products
XX
SQ Sequence 529 AA;

Alignment Scores:

Pred. No.:	5.75e-172	Length:	529
Score:	2738.00	Matches:	527
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	88.3%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x AAW85055 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529

RESULT 3
AAU27550
ID   AAU27550 standard; protein; 529 AA.
XX
AC   AAU27550;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
DR   N-PSDB; AAS43749.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Claim 8; Fig 2; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. The present sequence represents corneodesmosin amino acid
CC   sequence, as described in the invention
XX
SQ   Sequence 529 AA;

Alignment Scores:

Pred. No.:              5.75e-172       Length:         529
Score:                  2738.00         Matches:        527
Percent Similarity:     99.8%           Conservative:   1
Best Local Similarity:  99.6%           Mismatches:     1
Query Match:            88.3%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x AAU27550 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320
```

```
Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

```
RESULT 4
ABJ39148
ID   ABJ39148 standard; protein; 529 AA.
XX
AC   ABJ39148;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human corneodesmosin (CDSN) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 28
FT                   /note= "Encoded by CCAG"
FT   misc_difference 408
FT                   /note= "Encoded by TCC"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   N-PSDB; ABT43849, ABT43850.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
XX
SQ   Sequence 529 AA;
```

```
Alignment Scores:
Pred. No.:              5.75e-172      Length:          529
Score:                  2738.00        Matches:         527
Percent Similarity:     99.8%          Conservative:    1
Best Local Similarity:  99.6%          Mismatches:      1
Query Match:            88.3%          Indels:          0
DB:                     7              Gaps:            0

US-10-677-865A-2 (1-1590) x ABJ39148 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 5
ADN04226
ID ADN04226 standard; protein; 529 AA.
XX
AC ADN04226;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic protein sequence #308.
XX
KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR N-PSDB; ADN04225.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX
SQ Sequence 529 AA;

Alignment Scores:

Pred. No.:	5.75e-172	Length:	529
Score:	2738.00	Matches:	527
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	88.3%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADN04226 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 6
ADP12525
ID ADP12525 standard; protein; 529 AA.
XX
AC ADP12525;
XX
DT 12-AUG-2004 (first entry)
XX
DE Protein encoded by mRNA of the invention #135.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX

SQ Sequence 529 AA;

Alignment Scores:

Pred. No.:	9.07e-172	Length:	529
Score:	2735.00	Matches:	526
Percent Similarity:	99.8%	Conservative:	2
Best Local Similarity:	99.4%	Mismatches:	1
Query Match:	88.2%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADP12525 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerSerGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 7
AAY44989
ID AAY44989 standard; protein; 529 AA.
XX
AC AAY44989;
XX
DT 23-MAY-2000 (first entry)
XX
DE Human epidermal protein-6.
XX
KW Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW dermatological; anitarteriosclerotic; antiinflammatory;
KW immunosuppressive.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Peptide	1. .32
FT		/label= Signal_peptide
FT	Modified-site	3
FT		/note= "Potential phosphorylation site"
FT	Modified-site	37
FT		/note= "Potential phosphorylation site"

```
FT   Modified-site   50
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   57
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   80
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   172
FT                   /note= "N-glycosylated"
FT   Modified-site   182
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   287
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   395
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   440
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   N-PSDB; AAZ50582.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 1; Fig 4; 82pp; English.
XX
CC   The present sequence is human epidermal protein-6 (HEPI) expressed in
CC   breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC   cDNA library. HEPI-6 shares 99% identity with human S protein.
CC   Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC   for protein expression. The HEPI proteins are useful for the treatment of
CC   epithelial disorders, including dyshidrotic eczema and scabies, cell
CC   proliferative disorders including actinic keratosis and arteriosclerosis,
CC   and autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              8.84e-171      Length:         529
Score:                  2720.00        Matches:        523
Percent Similarity:     99.4%          Conservative:   3
Best Local Similarity:  98.9%          Mismatches:     3
Query Match:            87.7%          Indels:         0
DB:                     3              Gaps:           0

US-10-677-865A-2 (1-1590) x AAY44989 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetLeuAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrPheGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380
```

```
Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||:::|||   |||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerSerGlyLeuProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 8
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001 (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX

FH	Key	Location/Qualifiers
FT	Misc-difference	143
FT		/label= Ser, Asn
FT	Misc-difference	151
FT		/label= Ser, Cys
FT	Misc-difference	152
FT		/label= Ser, Arg
FT	Misc-difference	200
FT		/label= Ser, Phe
FT	Misc-difference	282
FT		/label= Pro, Thr
FT	Misc-difference	401
FT		/label= Ser, Gly
FT	Misc-difference	408
FT		/label= Ala, Ser
FT	Misc-difference	409
FT		/label= Gly, Val
FT	Misc-difference	410
FT		/label= Ser, Leu
FT	Misc-difference	453
FT		/label= Ser, Asn
FT	Misc-difference	527
FT		/label= Asp, Asn

XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H, Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

Alignment Scores:

Pred. No.:	2.43e-168	Length:	529
Score:	2683.00	Matches:	516
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	97.5%	Mismatches:	12
Query Match:	86.5%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x AAG64152 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||   ||||||||||||||||||||      ||||||||||||||||||||||||||
Db        141 SerGly***SerGlySerHisSerGlySer******SerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180
```

```
Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 Thr***GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 Lys***CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
                 ||||||||||||||||||          ||||||||||||||||||||||||||||||
Db        401 ***SerSerIleSerSerSer*********ProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSer***SerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||   ||||||
Db        521 ProGlnGlyGluLeuLeu***SerPro 529
```

RESULT 9
ABJ39147
ID ABJ39147 standard; protein; 486 AA.
XX
AC ABJ39147;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human S protein encoded by the skin corneodesmosin (CDSN) gene (S-gene).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW S protein; skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43848.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polypeptide sequence is the human S protein (corneodesmosin protein only
CC expressed in the skin), the target of the invention
XX
SQ Sequence 486 AA;

Alignment Scores:

Pred. No.:	1.08e-155	Length:	486
Score:	2491.00	Matches:	480
Percent Similarity:	99.8%	Conservative:	3
Best Local Similarity:	99.2%	Mismatches:	1
Query Match:	80.3%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39147 (1-486)

```
Qy         49 ATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuAlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGly 20

Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        21 ThrPheSerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 40

Qy       169 ACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC 228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySer 60

Qy       229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 SerIleSerSerAlaArgSerSerGlyGlySerSerGlySerSerSerGlySerSer 80

Qy       289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 IleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSer 100

Qy       349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGC 408
             ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db       101 TyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySer 120

Qy       409 AGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGC 468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 SerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSer 140

Qy       469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 SerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeu 160

Qy       529 CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCT 588
             ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       161 ProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSer 180

Qy       589 TCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGT 648
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 SerSerSerGlnThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArg 200

Qy       649 CCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 ProCysSerSerAspIleProAspSerProCysSerGlyGlyProIleValSerHisSer 220

Qy       709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 GlyProTyrIleProSerSerHisSerValSerGlyGlyGlnArgProValValValVal 240

Qy       769 GTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGT 828
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ValAspGlnHisGlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGly 260

Qy       829 GGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 GlyLeuProGlyLysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyr 280

Qy       889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGT 948
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 GluValValGlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGly 300

Qy       949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 LysIleTyrProValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyVal 320

Qy      1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 340

Qy      1069 ATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 IleProSerGlnSerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGly 360

Qy      1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSer 380

Qy      1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGC 1248
             ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db       381 ArgValProSerSerSerSerIleSerSerSerSerGlySerProTyrHisProCysGly 400

Qy      1249 AGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGT 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSerSerSer 420

Qy      1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 SerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGly 440

Qy      1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 HisProCysMetSerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHis 460

Qy      1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 ProAspProSerAlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCys 480

Qy      1489 CGCTCCATCCGG 1500
             ||||||||||||
Db       481 ArgSerIleArg 484
```

RESULT 10
ADQ21504
ID ADQ21504 standard; protein; 486 AA.
XX
AC ADQ21504;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4324.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4324; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 486 AA;

Alignment Scores:

Pred. No.:	1.08e-155	Length:	486
Score:	2491.00	Matches:	480
Percent Similarity:	99.8%	Conservative:	3
Best Local Similarity:	99.2%	Mismatches:	1
Query Match:	80.3%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADQ21504 (1-486)

```
Qy         49 ATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuAlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGly 20

Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrPheSerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 40

Qy        169 ACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySer 60

Qy        229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSer 80

Qy        289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSer 100

Qy        349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGC 408
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        101 TyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySer 120

Qy        409 AGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSer 140

Qy        469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeu 160

Qy        529 CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCT 588
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        161 ProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSer 180

Qy        589 TCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGT 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SerSerSerGlnThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArg 200

Qy        649 CCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProCysSerSerAspIleProAspSerProCysSerGlyGlyProIleValSerHisSer 220

Qy        709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProTyrIleProSerSerHisSerValSerGlyGlyGlnArgProValValValVal 240

Qy        769 GTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ValAspGlnHisGlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGly 260

Qy        829 GGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyLeuProGlyLysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyr 280

Qy        889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGT 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GluValValGlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGly 300

Qy        949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LysIleTyrProValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyVal 320

Qy       1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 340

Qy       1069 ATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleProSerGlnSerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGly 360

Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSer 380

Qy       1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGC 1248
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        381 ArgValProSerSerSerSerIleSerSerSerSerGlySerProTyrHisProCysGly 400

Qy       1249 AGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSerSerSer 420

Qy       1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGly 440

Qy       1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 HisProCysMetSerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHis 460

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 ProAspProSerAlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCys 480

Qy       1489 CGCTCCATCCGG 1500
              ||||||||||||
Db        481 ArgSerIleArg 484
```

RESULT 11
ABO60412
ID ABO60412 standard; protein; 219 AA.
XX
AC ABO60412;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #6646.
XX
KW Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 45; SEQ ID NO 34046; 80pp; English.

```
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe protein of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 219 AA;

Alignment Scores:
Pred. No.:              1.5e-63       Length:          219
Score:                  1092.00       Matches:         217
Percent Similarity:     99.5%         Conservative:    1
Best Local Similarity:  99.1%         Mismatches:      1
Query Match:            35.2%         Indels:          0
DB:                     8             Gaps:            0

US-10-677-865A-2 (1-1590) x ABO60412 (1-219)

Qy         88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ThrLeuAlaLysSerIleGlyThrPheSerAspProCysLysAspProThrArgIleThr 20

Qy        148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSer 40

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlySerSerSerSerGlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSer 60

Qy        268 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGly 80

Qy        328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCC 387
              ||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db         81 ThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySerSerLeuGlnGlyAlaSer 100

Qy        388 GGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerSerGlnLeuGlySerSerSerSerHisSerGlySerSerGlySerHisSerGly 120

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerSerSerSerHisSerSerSerSerSerSerPheGlnPheSerSerSerSerPheGln 140

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db        141 ValGlyAsnGlySerAlaLeuProThrAsnAspAsnSerTyrArgGlyIleLeuAsnPro 160

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 SerGlnProGlyGlnSerSerSerSerSerGlnThrSerGlyValSerSerSerGlyGln 180

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SerValSerSerAsnGlnArgProCysSerSerAspIleProAspSerProCysSerGly 200

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyProIleValSerHisSerGlyProTyrIleProSerSerHisSerValSerGly 219

RESULT 12
ABJ39195
ID   ABJ39195 standard; protein; 146 AA.
XX
AC   ABJ39195;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment c, cleavage product of human S (corneodesmosin skin) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
```

```
CC   the invention
XX
SQ   Sequence 146 AA;

Alignment Scores:
Pred. No.:              6.05e-38        Length:         146
Score:                  703.00          Matches:        144
Percent Similarity:     99.3%           Conservative:   1
Best Local Similarity:  98.6%           Mismatches:     1
Query Match:            22.7%           Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x ABJ39195 (1-146)

Qy        163 TGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCT 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSer 20

Qy        223 GGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGA 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGly 40

Qy        283 TCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAG 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGln 60

Qy        343 GTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTG 402
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db         61 ValSerTyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeu 80

Qy        403 GGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCAT 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GlySerSerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHis 100

Qy        463 TCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerSerSerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySer 120

Qy        523 GCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAA 582
              |||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        121 AlaLeuProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGln 140

Qy        583 AGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||
Db        141 SerSerSerSerSerGln 146

RESULT 13
ABJ39196
ID   ABJ39196 standard; protein; 111 AA.
XX
AC   ABJ39196;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment d, glycine rich sticky domain 1 of the human S protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 111 AA;

Alignment Scores:
Pred. No.:              1.45e-26        Length:         111
Score:                  530.00          Matches:        110
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.1%           Mismatches:     0
Query Match:            17.1%           Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x ABJ39196 (1-111)

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSerAlaArgSerSer 20

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGly 40

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         41 SerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySerSer 60

Qy        373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGC 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHisSerGlySerSer 80

Qy        433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSerPheGlnPheSer 100

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCT 525
              |||||||||||||||||||||||||||||||||
Db        101 SerSerSerPheGlnValGlyAsnGlySerAla 111

RESULT 14
ABJ39197
ID   ABJ39197 standard; protein; 76 AA.
XX
AC   ABJ39197;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment e, glycine rich sticky domain 2 of the human S protein.
```

XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 76 AA;

Alignment Scores:

Pred. No.:	4.98e-18	Length:	76
Score:	400.00	Matches:	76
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	12.9%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39197 (1-76)

```
Qy       1123 GGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCC 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyGlyValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSer 20

Qy       1183 AGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCC 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerSerArgValProSerSerSerSerIleSerSerSerAlaGlySerProTyrHisPro 40

Qy       1243 TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 CysGlySerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSer 60

Qy       1303 TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerSerSerGlnSerSerGlyLysIleIleLeuGlnProCysGlySer 76
```

RESULT 15
AAU91255
ID AAU91255 standard; protein; 71 AA.
XX
AC AAU91255;
XX
DT 18-JUN-2002 (first entry)
XX
DE Nucleic acid selection related corneodesmosin peptide.
XX
KW Nucleic acid selection; viral genome; sindbis virus.
XX
OS Mus musculus.
XX
PN WO200216572-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-CA001169.
XX
PR 18-AUG-2000; 2000US-00641931.
XX
PA (PHEN-) PHENOGENE THERAPEUTIQUES INC.
XX
PI Lanctot C, Moffat P, Salois P;
XX
DR WPI; 2002-257913/30.
XX
PT Selecting exogenous nucleic acids having a desired feature, particularly
PT from a library of nucleic acids, involves using a viral genome.
XX
PS Example 4; Page 55; 88pp; English.
XX
CC The invention describes the use of a viral genome for selecting an
CC exogenous nucleic acid having a desired feature. The method requires a
CC viral genome, e.g. sindbis virus, capable of expressing the exogenous
CC nucleic acid in a host cell and selecting among these the ones having the
CC desired feature or predefined function. The method removes the expensive
CC and time-consuming task of selecting cells that express a gene of
CC interest. The method is more rapid, efficient and accurate for selecting
CC a particular nucleic acid having a desired feature, characteristic or
CC function. This sequence represents a peptide encoded by a nucleic acid
CC isolated using the method described in the invention
XX
SQ Sequence 71 AA;

Alignment Scores:

Pred. No.:	5.42e-16	Length:	71
Score:	369.00	Matches:	69
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAU91255 (1-71)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGT 207
              |||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSer 69
```

Search completed: March 21, 2006, 21:58:41
Job time : 166 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:07:39 ; Search time 27.3 Seconds
(without alignments)
4867.023 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152027_26248/app_query.fasta_1
-DB=Published_Applications_AA_Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_134_@runat_21032006_152027_26248 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2738	88.3	529	4	US-10-316-175-15	Sequence 15, Appl
2	2738	88.3	529	5	US-10-677-865-1	Sequence 1, Appli
3	2720	87.7	529	5	US-10-883-805-6	Sequence 6, Appli
4	2693	86.8	529	4	US-10-204-884-3	Sequence 3, Appli
5	2683	86.5	529	4	US-10-164-230-61	Sequence 61, Appl
6	2491	80.3	486	4	US-10-316-175-12	Sequence 12, Appl
7	2491	80.3	486	5	US-10-723-860-4324	Sequence 4324, Ap
8	2491	80.3	486	5	US-10-883-805-16	Sequence 16, Appl
9	1092	35.2	219	4	US-10-029-386-34046	Sequence 34046, A
10	703	22.7	146	4	US-10-316-175-71	Sequence 71, Appl
11	530	17.1	111	4	US-10-316-175-72	Sequence 72, Appl
12	421	13.6	81	4	US-10-316-175-73	Sequence 73, Appl
13	335.5	10.8	950	6	US-11-097-143-41331	Sequence 41331, A
14	315	10.2	2112	6	US-11-097-143-8001	Sequence 8001, Ap
15	310	10.0	58	4	US-10-316-175-70	Sequence 70, Appl
16	298	9.6	55	4	US-10-316-175-69	Sequence 69, Appl
17	292.5	9.4	528	3	US-09-840-746-20	Sequence 20, Appl
18	290	9.3	668	4	US-10-437-963-134726	Sequence 134726,
19	290	9.3	2280	6	US-11-097-143-11742	Sequence 11742, A
20	285	9.2	697	4	US-10-425-114-41545	Sequence 41545, A
21	275.5	8.9	708	4	US-10-425-115-313976	Sequence 313976,
22	270.5	8.7	496	4	US-10-437-963-187970	Sequence 187970,
23	265.5	8.6	842	6	US-11-097-143-26685	Sequence 26685, A
24	265.5	8.6	864	6	US-11-097-143-40749	Sequence 40749, A
25	263	8.5	258	4	US-10-104-047-3034	Sequence 3034, Ap
26	263	8.5	566	4	US-10-425-115-336849	Sequence 336849,
27	261.5	8.4	1367	3	US-09-801-368-108	Sequence 108, App
28	260	8.4	465	4	US-10-437-963-174113	Sequence 174113,
29	259	8.3	1459	4	US-10-369-493-6418	Sequence 6418, Ap
30	257	8.3	384	4	US-10-479-638-19	Sequence 19, Appl
31	256.5	8.3	693	4	US-10-108-260A-3849	Sequence 3849, Ap
32	253.5	8.2	1480	4	US-10-408-765A-2239	Sequence 2239, Ap
33	253.5	8.2	1828	4	US-10-221-625-15	Sequence 15, Appl
34	253	8.2	598	4	US-10-282-122A-64954	Sequence 64954, A
35	253	8.2	903	4	US-10-156-761-11093	Sequence 11093, A
36	252	8.1	2586	6	US-11-097-143-27426	Sequence 27426, A
37	251.5	8.1	2971	4	US-10-146-473-50	Sequence 50, Appl
38	251.5	8.1	2971	5	US-10-732-923-8849	Sequence 8849, Ap
39	250	8.1	1357	6	US-11-097-143-39117	Sequence 39117, A
40	249.5	8.0	1217	4	US-10-311-406-2	Sequence 2, Appli
41	249	8.0	437	4	US-10-425-115-215389	Sequence 215389,
42	248	8.0	1455	6	US-11-097-143-15912	Sequence 15912, A
43	247.5	8.0	1566	4	US-10-177-293-366	Sequence 366, App
44	247.5	8.0	1566	4	US-10-435-696-30	Sequence 30, Appl
45	247.5	8.0	1566	4	US-10-425-970-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-15

Alignment Scores:

Pred. No.:	1.12e-150	Length:	529
Score:	2738.00	Matches:	527
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	88.3%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-316-175-15 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 2
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/10/677,865
; CURRENT FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: US/09/381,656
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-10-677-865-1

Alignment Scores:

Pred. No.:	1.12e-150	Length:	529
Score:	2738.00	Matches:	527
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	88.3%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-677-865-1 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCCTGCCTCACTGGGAAGGGT 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
                |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
                ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
                |||||||||||||||||||||||||
Db         521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 3
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: LAL, Preeti
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
; FILE REFERENCE: PF-0567 PCT
; CURRENT APPLICATION NUMBER: US/10/883,805
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 09/123,531
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 09/206,817
; PRIOR FILING DATE: 1998-12-07
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

Alignment Scores:

Pred. No.:	1.24e-149	Length:	529
Score:	2720.00	Matches:	523

```
Percent Similarity:     99.4%          Conservative:  3
Best Local Similarity:  98.9%          Mismatches:    3
Query Match:            87.7%          Indels:        0
DB:                     5              Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-883-805-6 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetLeuAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ThrPheGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||:::|||   ||||||||||||||||||||||||||||||
Db        401 SerSerSerIleSerSerSerSerGlyLeuProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529

RESULT 4
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
;   FEATURE:
```

```
;   NAME/KEY: SITE
;   LOCATION: (143)..(143)
;   OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (401)..(401)
;   OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (409)..(409)
;   OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (410)..(410)
;   OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (527)..(527)
;   OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

Alignment Scores:
Pred. No.:              4.56e-148      Length:          529
Score:                  2693.00        Matches:         518
Percent Similarity:     98.1%          Conservative:    1
Best Local Similarity:  97.9%          Mismatches:      10
Query Match:            86.8%          Indels:          0
DB:                     4              Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-204-884-3 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||
```

```
RESULT 5
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

Alignment Scores:
Pred. No.:              1.73e-147       Length:         529
Score:                  2683.00         Matches:        516
Percent Similarity:     97.7%           Conservative:   1
Best Local Similarity:  97.5%           Mismatches:     12

Query Match:            86.5%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-164-230-61 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||  ||||||||||||||||||||      ||||||||||||||||||||||||||
Db        141 SerGly***SerGlySerHisSerGlySer******SerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 Thr***GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 Lys***CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340
```

```
Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
                ||||||||||||||||||          ||||||||||||||||||||||||||||||
Db        401 ***SerSerIleSerSerSer*********ProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSer***SerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
             |||||||||||||||||   ||||||
Db        521 ProGlnGlyGluLeuLeu***SerPro 529

RESULT 6
US-10-316-175-12
; Sequence 12, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-12

Alignment Scores:
Pred. No.:              2.37e-136       Length:         486
Score:                  2491.00         Matches:        480
Percent Similarity:     99.8%           Conservative:   3
Best Local Similarity:  99.2%           Mismatches:     1
Query Match:            80.3%           Indels:         0
DB:                     4               Gaps:           0
```

```
US-10-677-865A-2 (1-1590) x US-10-316-175-12 (1-486)

Qy         49 ATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuAlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGly 20

Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrPheSerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 40

Qy        169 ACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySer 60

Qy        229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSer 80

Qy        289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSer 100

Qy        349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGC 408
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        101 TyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySer 120

Qy        409 AGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSer 140

Qy        469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeu 160

Qy        529 CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCT 588
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        161 ProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSer 180

Qy        589 TCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGT 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SerSerSerGlnThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArg 200

Qy        649 CCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProCysSerSerAspIleProAspSerProCysSerGlyGlyProIleValSerHisSer 220

Qy        709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProTyrIleProSerSerHisSerValSerGlyGlyGlnArgProValValValVal 240

Qy        769 GTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ValAspGlnHisGlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGly 260

Qy        829 GGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyLeuProGlyLysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyr 280

Qy        889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGT 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GluValValGlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGly 300

Qy        949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LysIleTyrProValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyVal 320

Qy       1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 340

Qy       1069 ATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleProSerGlnSerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGly 360
```

```
Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSer 380

Qy       1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGC 1248
             ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db        381 ArgValProSerSerSerSerIleSerSerSerSerGlySerProTyrHisProCysGly 400

Qy       1249 AGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGT 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSerSerSer 420

Qy       1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGly 440

Qy       1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 HisProCysMetSerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHis 460

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 ProAspProSerAlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCys 480

Qy       1489 CGCTCCATCCGG 1500
             ||||||||||||
Db        481 ArgSerIleArg 484

RESULT 7
US-10-723-860-4324
; Sequence 4324, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4324
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-4324

Alignment Scores:
Pred. No.:              2.37e-136       Length:         486
Score:                  2491.00         Matches:        480
Percent Similarity:     99.8%           Conservative:   3
Best Local Similarity:  99.2%           Mismatches:     1
Query Match:            80.3%           Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-723-860-4324 (1-486)

Qy         49 ATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuAlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGly 20

Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrPheSerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 40

Qy        169 ACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySer 60

Qy        229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSer 80

Qy        289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSer 100

Qy        349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGC 408
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        101 TyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySer 120

Qy        409 AGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSer 140

Qy        469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeu 160

Qy        529 CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCT 588
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        161 ProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSer 180

Qy        589 TCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGT 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SerSerSerGlnThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArg 200

Qy        649 CCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProCysSerSerAspIleProAspSerProCysSerGlyGlyProIleValSerHisSer 220

Qy        709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProTyrIleProSerSerHisSerValSerGlyGlyGlnArgProValValValVal 240

Qy        769 GTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ValAspGlnHisGlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGly 260

Qy        829 GGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyLeuProGlyLysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyr 280

Qy        889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGT 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GluValValGlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGly 300

Qy        949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LysIleTyrProValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyVal 320

Qy       1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 340

Qy       1069 ATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleProSerGlnSerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGly 360

Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSer 380

Qy       1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGC 1248
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db        381 ArgValProSerSerSerSerIleSerSerSerSerGlySerProTyrHisProCysGly 400

Qy       1249 AGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSerSerSer 420
```

```
Qy       1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGly 440

Qy       1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 HisProCysMetSerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHis 460

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 ProAspProSerAlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCys 480

Qy       1489 CGCTCCATCCGG 1500
             ||||||||||||
Db        481 ArgSerIleArg 484

RESULT 8
US-10-883-805-16
; Sequence 16, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 16
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No: g414810
US-10-883-805-16

Alignment Scores:
Pred. No.:              2.37e-136      Length:         486
Score:                  2491.00        Matches:        480
Percent Similarity:     99.8%          Conservative:   3
Best Local Similarity:  99.2%          Mismatches:     1
Query Match:            80.3%          Indels:         0
DB:                     5              Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-883-805-16 (1-486)

Qy         49 ATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuAlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGly 20

Qy        109 ACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ThrPheSerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 40

Qy        169 ACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySer 60

Qy        229 TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSer 80

Qy        289 ATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGC 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 IleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSer 100

Qy        349 TACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGC 408
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        101 TyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySer 120

Qy        409 AGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGC 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSer 140

Qy        469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 SerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeu 160

Qy        529 CCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCT 588
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        161 ProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSer 180

Qy        589 TCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGT 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SerSerSerGlnThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArg 200

Qy        649 CCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ProCysSerSerAspIleProAspSerProCysSerGlyGlyProIleValSerHisSer 220

Qy        709 GGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTG 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlyProTyrIleProSerSerHisSerValSerGlyGlyGlnArgProValValValVal 240

Qy        769 GTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ValAspGlnHisGlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGly 260

Qy        829 GGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTAC 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyLeuProGlyLysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyr 280

Qy        889 GAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGT 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GluValValGlyGlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGly 300

Qy        949 AAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LysIleTyrProValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyVal 320

Qy       1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 340

Qy       1069 ATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleProSerGlnSerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGly 360

Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSer 380

Qy       1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGC 1248
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        381 ArgValProSerSerSerSerIleSerSerSerSerGlySerProTyrHisProCysGly 400

Qy       1249 AGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSerSerSer 420
```

```
Qy       1309 TCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGT 1368
Db        421 SerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGly 440

Qy       1369 CACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCAT 1428
Db        441 HisProCysMetSerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHis 460

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
Db        461 ProAspProSerAlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCys 480

Qy       1489 CGCTCCATCCGG 1500
Db        481 ArgSerIleArg 484

RESULT 9
US-10-029-386-34046
; Sequence 34046, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34046
;   LENGTH: 219
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000510.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.21
;   OTHER INFORMATION: SWISSPROT HIT: Q15517, EVALUE 1.00e-24
US-10-029-386-34046

Alignment Scores:
Pred. No.:              3.21e-55      Length:         219
Score:                  1092.00       Matches:        217
Percent Similarity:     99.5%         Conservative:   1
Best Local Similarity:  99.1%         Mismatches:     1
Query Match:            35.2%         Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-029-386-34046 (1-219)

Qy         88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
Db          1 ThrLeuAlaLysSerIleGlyThrPheSerAspProCysLysAspProThrArgIleThr 20

Qy        148 TCCCCTAACGACCCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
Db         21 SerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSer 40

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 267
Db         41 GlySerSerSerSerGlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSer 60

Qy        268 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
Db         61 GlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGly 80

Qy        328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCC 387
Db         81 ThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySerSerLeuGlnGlyAlaSer 100

Qy        388 GGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA 447
Db        101 GlySerSerGlnLeuGlySerSerSerSerHisSerGlySerSerGlySerHisSerGly 120

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507
Db        121 SerSerSerSerHisSerSerSerSerSerSerPheGlnPheSerSerSerSerPheGln 140

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567
Db        141 ValGlyAsnGlySerAlaLeuProThrAsnAspAsnSerTyrArgGlyIleLeuAsnPro 160

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627
Db        161 SerGlnProGlyGlnSerSerSerSerSerGlnThrSerGlyValSerSerSerGlyGln 180

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687
Db        181 SerValSerSerAsnGlnArgProCysSerSerAspIleProAspSerProCysSerGly 200

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
Db        201 GlyProIleValSerHisSerGlyProTyrIleProSerSerHisSerValSerGly 219

RESULT 10
US-10-316-175-71
; Sequence 71, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-71

Alignment Scores:
Pred. No.:              1.15e-32      Length:         146
Score:                  703.00        Matches:        144
Percent Similarity:     99.3%         Conservative:   1
Best Local Similarity:  98.6%         Mismatches:     1
Query Match:            22.7%         Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-71 (1-146)

Qy        163 TGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCT 222
Db          1 CysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSer 20

Qy        223 GGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGA 282
Db         21 GlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGly 40

Qy        283 TCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAG 342
Db         41 SerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGln 60

Qy        343 GTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTG 402
Db         61 ValSerTyrSerSerGlySerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeu 80
```

```
Qy        403 GGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCAT 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GlySerSerSerSerHisSerGlySerSerGlySerHisSerGlySerSerSerSerHis 100

Qy        463 TCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerSerSerSerSerSerPheGlnPheSerSerSerSerPheGlnValGlyAsnGlySer 120

Qy        523 GCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAA 582
              |||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        121 AlaLeuProThrAsnAspAsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGln 140

Qy        583 AGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||
Db        141 SerSerSerSerSerGln 146

RESULT 11
US-10-316-175-72
; Sequence 72, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72
;   LENGTH: 111
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-72

Alignment Scores:
Pred. No.:              1.23e-22      Length:          111
Score:                  530.00        Matches:         110
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  99.1%         Mismatches:      0
Query Match:            17.1%         Indels:          0
DB:                     4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-316-175-72 (1-111)

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSerAlaArgSerSer 20

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGly 40

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         41 SerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySerSer 60

Qy        373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGC 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHisSerGlySerSer 80

Qy        433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSerPheGlnPheSer 100

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCT 525
              |||||||||||||||||||||||||||||||||
Db        101 SerSerSerPheGlnValGlyAsnGlySerAla 111

RESULT 12
US-10-316-175-73
; Sequence 73, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-73

Alignment Scores:
Pred. No.:              2.54e-16      Length:          81
Score:                  421.00        Matches:         81
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            13.6%         Indels:          0
DB:                     4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-316-175-73 (1-81)

Qy       1123 GGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCC 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyGlyValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSer 20

Qy       1183 AGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCC 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerSerArgValProSerSerSerSerIleSerSerSerAlaGlySerProTyrHisPro 40

Qy       1243 TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 CysGlySerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSer 60

Qy       1303 TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerSerSerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSer 80

Qy       1363 TCT 1365
              |||
Db         81 Ser 81

RESULT 13
US-11-097-143-41331
; Sequence 41331, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
```

```
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41331
;   LENGTH: 950
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-41331

Alignment Scores:
Pred. No.:              2.66e-11        Length:         950
Score:                  335.50          Matches:        170
Percent Similarity:     39.0%           Conservative:   32
Best Local Similarity:  32.8%           Mismatches:     174
Query Match:            10.8%           Indels:         142
DB:                     6               Gaps:           28

US-10-677-865A-2 (1-1590) x US-11-097-143-41331 (1-950)

Qy         94 GCTAAGAGCATTGGCACCTTCTCAGAC---CCCTGTAAGGACCCCACGCGTATCACCTCC 150
Db        494 AlaProSerValSerSerPheValProLeuProSerAlaProSerThrAsnTyrGlyAla 513

Qy        151 CCTAAC---GACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGC-------------- 192
Db        514 ProSerLysThrGlnSerLeuGlySerSerGlyTyrSerSerGlyProSerSerSerTyr 533

Qy        193 --------------------TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231
Db        534 GluAlaProValAlaProProSerSerSerTyrGlyAlaProSerSerSerPheGlnPro 553

Qy        232 ATTTCCAGTGCCAGAAGCTCTGGTGGT---------GGCTCCAGTGGTAGCTCCAGCGGA 282
Db        554 IleSerProProSerSerSerTyrGlyAlaProSerSerGlySerGlySerSerSerGly 573

Qy        283 TCCAGCATTGCC----------------------------CAGGGTGGTTCTGCAGGA 312
Db        574 SerPheSerAlaAlaProSerSerLeuTyrSerAlaProSerLysGlySerSerGlyGly 593

Qy        313 TCTTTTAAG------CCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGC 366
Db        594 SerPheGlnSerAlaProSerSerSerTyrSerAlaProSerAlaSerAlaAsnSerGly 613

Qy        367 TCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGA 426
Db        614 GlySerTyrProSerAlaProSerSerSerTyrSerAlaProSerSerSerSerSerSer 633

Qy        427 AGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAG 486
Db        634 GlyGlyProTyrAlaSerAlaProSerSerSerTyrSerAlaProSerSerGlySerAsn 653

Qy        487 TTCAGC--------------------AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCT 525
Db        654 SerGlyGlyProTyrProAlaAlaProSerSerSerTyrSerAlaProSerAlaSerAla 673

Qy        526 CTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGC 585
Db        674 --------------AsnSer------GlyGlySerTyrProSerAlaPro---SerSer 685

Qy        586 TCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAG 645
```

```
Db        686 SerTyrSerAlaProSerProGlySerAsnSerGlyGlyProTyrProAlaAla------ 703

Qy        646 CGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGC------AGTGGAGGGCCCATCGTC 699
Db        704 ---ProSerSerSerTyrSerAlaProSerProSerAlaAsnSerGlyGlyProTyrAla 722

Qy        700 TCG------CACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGG 753
Db        723 SerAlaProSerSerSerTyrSerAlaProSerSerSerSerAsnSerGlyGlyProTyr 742

Qy        754 CCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCC 813
Db        743 ----------------------------AlaAlaAlaProSerSerSerTyrSerAla 752

Qy        814 CCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAA 873
Db        753 ProSerSerSerSerSer---SerGlyGlyProTyrPro-------------------- 764

Qy        874 TCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATG 933
Db        765 ----------------------SerAlaProSerSerSerTyrSerAlaProSerSer 776

Qy        934 ACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAA 993
Db        777 SerLeuSerSerGlyGlyProTyrPro---------------------------------- 785

Qy        994 GGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAG------GGCAAA 1047
Db        786 ---SerAlaProSerSerSerTyrAlaAlaProSerProSerSerAsnSerGlyGlyPro 804

Qy       1048 TACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTC 1107
Db        805 Tyr------------ProAlaAlaProSerAsnSer------TyrSerAlaProIleAla 818

Qy       1108 CAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGA 1167
Db        819 ProProSerSerSerTyrGly------------------------------------- 825

Qy       1168 CCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGT 1227
Db        826 ------AlaProAlaSerGlyProSerProSerPheSerAlaProSerSerSerTyrGly 843

Qy       1228 TCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGC 1287
Db        844 AlaPro------SerThrGlySer--------------------------GlySerSer 852

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGC 1347
Db        853 SerPheSerSerSerSerSerSerPheSerGly------------------------Ala 864

Qy       1348 AGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGG 1407
Db        865 SerSerSerSerSerAlaGlyTyrProSerAlaProSerSerSerTyr---------Gly 881

Qy       1408 GGCCCCGAT---GGCTCTCCCCAT--------------CCTGATCCCTCCGCTGGTGCC 1449
Db        882 AlaProSerThrGlySerGlyHisSerPheSerSerAlaProSerSerSerTyrSerAla 901

Qy       1450 AAGCCCTGT---GGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCC 1498
Db        902 ProProAlaGlyGlySerSerSerSerGlyProTyr-ProSerAlaProSer 918

RESULT 14
US-11-097-143-8001
; Sequence 8001, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
```

```
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8001
;   LENGTH: 2112
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-8001

Alignment Scores:
Pred. No.:                4.3e-10        Length:          2112
Score:                    315.00         Matches:         152
Percent Similarity:       39.8%          Conservative:    50
Best Local Similarity:    29.9%          Mismatches:      192
Query Match:              10.2%          Indels:          114
DB:                       6              Gaps:            21

US-10-677-865A-2 (1-1590) x US-11-097-143-8001 (1-2112)

Qy        100 AGCATTGGCACCTTCTCAGACCCC-----------TGTAAGGACCCCACGCGTATCACC 147
Db        733 SerProGlyThrValTrpAspProAspThrLysGlyCysAsnHisProThrAspValGln 752

Qy        148 TCCCCTAACGACCCCTGCCTC---------ACTGGGAAGGGTGACTCCAGCGGCTTCAGT 198
Db        753 LysGluGlnCysLysAlaMetAlaAsnGlySerGlySerSerSerSerGlnGlySerSer 772

Qy        199 AGCTACAGTGGCTCCAGCAGTTCT---GGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGT 255
Db        773 SerAsnGlnGlySerSerSerSerGlnGlySerSerSerAsnGlnGlySerSerSerAsn 792

Qy        256 GGTGGCTCCAGT---------GGTAGCTCCAGC------GGATCCAGCATTGCCCAGGGT 300
Db        793 GlnGlySerSerSerAsnGlnGlySerSerSerAsnGlnGlySerSerSerAsnGlnGly 812

Qy        301 GGTTCTGCA---------GGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTAC 351
Db        813 SerSerSerAsnGlnGlySerSerSerAsnGlnGlySerSerSerAsnGlnGlySerSer 832

Qy        352 TCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCC---------GGTTCCTCCCAGCTG 402
Db        833 SerAsnGlnGlySerSerSerSerGlnGlySerSerSerAsnGlnGlySerSerSerAsn 852

Qy        403 GGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCAT 462
Db        853 GlnGlySerSerSerAsnGlnGlySerSer---SerAsnGlnGlySerSerSerAsnGln 871

Qy        463 TCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCT 522
Db        872 GlySerSerSerSerGlnGlySerSerSerAsnGluGlySerSerSerAsnGlnGlySer 891

Qy        523 GCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAA 582
Db        892 SerSerSerGlnGlySerSerSerAsnGlnGly-----------SerSerSerAsnGln 907

Qy        583 AGCTCTTCCTCTTCCCAAACCTCTGGGGTA---------TCCAGCAGTGGCCAAAGCGTC 633

Db        908 GlySerSerSerAsnGlnGlySerSerSerAsnGlnGlySerSerSerAsnGlnGlySer 927

Qy        634 AGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCC 693
Db        928 SerSerAsnGlnGlySerSerSerAsnGlnSerSerSerSerAsnGlnSer--------- 944

Qy        694 ATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGG 753
Db        945 -----------SerSerAsnGlnSerSerSerAsnGlnSerSerSerSerAsnGln--- 959

Qy        754 CCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCC 813
Db        959 ------------------------------------------------------------ 959

Qy        814 CCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAA 873
Db        960 ------ThrSerSerSerThrThrGlnLysProPheLysProAlaGluLysCysGlu--- 976

Qy        874 TCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATG 933
Db        977 ----------------------------SerGluGluThrPheLeuAlaAspAsnGlu 986

Qy        934 ACCTACAGTAAG--------------------GGTAAAATCTATCCTGTGGGCTACTTC 972
Db        987 AsnCysSerLysPheTyrArgCysValAspAsnGlyLys-----------GlyGlyPhe 1002

Qy        973 ACCAAAGAG----------------------------AACCCTGTGAAAGGCTCTCCA 1002
Db       1003 ThrLysValSerPheThrCysProProAsnThrLeuTrpAspProGluAlaAsnSerCys 1022

Qy       1003 GGGGTCCCTTCCTTTGCAGCTGGGCCCCCC-----------------ATCTCTGAGGGC 1044
Db       1023 AsnHisProAspGlnIleGlnThrLysProLeuLysCysLysLysValValSerGlnGly 1042

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
Db       1043 Gly---SerSerSerAsnSerThrSerAsnSerSerSerSerSerAsnAsn--------- 1058

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
Db       1059 --------------SerGlySerSerSerAsnSerGlySerSerSerSerSerSerSer 1073

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
Db       1074 SerAsnSerGlySerSerSerAsnThrGlySerSerSerAsnSerGlyAlaSerSerSer 1093

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
Db       1094 Gly-----------------GlySerSerAsnGlnGlySerSerSerAsnSerGlySer 1107

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
Db       1108 SerSerGlySerAsnSerSerGlyAsnGlnSerThrSer-----------SerSerThr 1123

Qy       1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
Db       1124 SerSerSerSerSerSerSerAsnAsnAsnAsnGlnGlySerSerSerSerSerSerSer 1143

Qy       1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
Db       1144 SerSerSerSerThrSerSerLysProAsnProSerGluThrCysLysValAsnGlyGln 1163

Qy       1465 AGCAGTGCTGGAAAGATCCCCTGC 1488
Db       1164 PheIleGlyAspArgSerAspCys 1171

RESULT 15
US-10-316-175-70
; Sequence 70, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
```

```
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-70

Alignment Scores:
Pred. No.:              6.86e-10        Length:         58
Score:                  310.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            10.0%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-70 (1-58)

Qy       1414 GATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspGlySerProHisProAspProSerAlaGlyAlaLysProCysGlySerSerSerAla 20

Qy       1474 GGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyLysIleProCysArgSerIleArgAspIleLeuAlaGlnValLysProLeuGlyPro 40

Qy       1534 CAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlnLeuAlaAspProGluValPheLeuProGlnGlyGluLeuLeuAspSerPro 58


Search completed: March 21, 2006, 22:13:57
Job time : 156.5 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:09:39 ; Search time 2.8 Seconds
(without alignments)
3250.742 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 339260

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152031_26308/app_query.fasta_1
-DB=Published_Applications_AA_New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_8_@runat_21032006_152031_26308 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	263	8.5	258	7	US-11-072-512-3034	Sequence 3034, Ap
	2	262.5	8.5	1274	7	US-11-096-568A-28959	Sequence 28959, A
	3	262.5	8.5	1480	7	US-11-096-568A-28958	Sequence 28958, A
	4	262	8.4	1267	7	US-11-096-568A-28960	Sequence 28960, A
	5	240	7.7	1377	6	US-10-821-234-1070	Sequence 1070, Ap
	6	239.5	7.7	1742	7	US-11-182-016-23	Sequence 23, Appl
	7	234.5	7.6	1970	6	US-10-821-234-1641	Sequence 1641, Ap
c	8	232	7.5	7465	7	US-11-087-099-7521	Sequence 7521, Ap
	9	228.5	7.4	413	7	US-11-096-568A-20771	Sequence 20771, A
c	10	223.5	7.2	783	7	US-11-052-554A-157	Sequence 157, App
	11	221	7.1	941	7	US-11-169-232-14	Sequence 14, Appl
	12	221	7.1	1022	7	US-11-169-232-84	Sequence 84, Appl
	13	221	7.1	1038	7	US-11-169-232-74	Sequence 74, Appl
	14	221	7.1	1049	7	US-11-169-232-58	Sequence 58, Appl
	15	221	7.1	1140	7	US-11-169-232-104	Sequence 104, App
	16	221	7.1	1270	7	US-11-169-232-44	Sequence 44, Appl
	17	221	7.1	1311	7	US-11-169-232-42	Sequence 42, Appl
	18	221	7.1	1313	7	US-11-169-232-142	Sequence 142, App
	19	221	7.1	1314	7	US-11-169-232-50	Sequence 50, Appl
	20	221	7.1	1320	7	US-11-169-232-46	Sequence 46, Appl
	21	221	7.1	1320	7	US-11-169-232-60	Sequence 60, Appl
	22	221	7.1	1354	7	US-11-169-232-48	Sequence 48, Appl
	23	221	7.1	1361	7	US-11-169-232-40	Sequence 40, Appl
	24	221	7.1	1363	7	US-11-169-232-52	Sequence 52, Appl
	25	221	7.1	1404	7	US-11-169-232-2	Sequence 2, Appli
	26	221	7.1	1404	7	US-11-169-232-62	Sequence 62, Appl
	27	219.5	7.1	596	6	US-10-063-703-100	Sequence 100, App
	28	219.5	7.1	596	7	US-11-102-240-100	Sequence 100, App
	29	219.5	7.1	596	7	US-11-103-195-100	Sequence 100, App
	30	218.5	7.0	1236	6	US-10-873-528-109	Sequence 109, App
	31	216	7.0	1560	7	US-11-059-982-1	Sequence 1, Appli
	32	215.5	6.9	5179	7	US-11-108-172-1068	Sequence 1068, Ap
	33	214	6.9	1381	7	US-11-052-554A-138	Sequence 138, App
	34	210.5	6.8	1306	7	US-11-052-554A-139	Sequence 139, App
c	35	210.5	6.8	1381	7	US-11-052-554A-138	Sequence 138, App
	36	210	6.8	440	6	US-10-063-703-52	Sequence 52, Appl
	37	210	6.8	440	7	US-11-102-240-52	Sequence 52, Appl
	38	210	6.8	440	7	US-11-103-195-52	Sequence 52, Appl
	39	208	6.7	1466	7	US-11-186-284-33	Sequence 33, Appl
c	40	207	6.7	749	7	US-11-052-554A-148	Sequence 148, App
	41	206	6.6	1464	7	US-11-000-463-243	Sequence 243, App
	42	204.5	6.6	1065	7	US-11-201-606-16	Sequence 16, Appl
	43	204	6.6	1166	6	US-10-821-234-964	Sequence 964, App
	44	203.5	6.6	957	7	US-11-108-172-1065	Sequence 1065, Ap
	45	202	6.5	1464	6	US-10-501-035-331	Sequence 331, App

ALIGNMENTS

RESULT 1
US-11-072-512-3034
; Sequence 3034, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3034
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3034

Alignment Scores:

Pred. No.:	1.84e-08	Length:	258
Score:	263.00	Matches:	94
Percent Similarity:	43.7%	Conservative:	20
Best Local Similarity:	36.0%	Mismatches:	113
Query Match:	8.5%	Indels:	34
DB:	7	Gaps:	5

US-10-677-865A-2 (1-1590) x US-11-072-512-3034 (1-258)

```
Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
              |||      |||:::     :::||||||      ||||||                |||
Db         15 SerThrLysCysArgCysGlySerArgIleAlaGlyProAsn---------AlaLeuGly 31

Qy        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCC---------------------- 213
                 |||   |||         ||||||   |||   |||
Db         32 SerGlyGlySerArgSerSerSerSerSerSerArgSerIleLeuSerSerSerIleLeu 51

Qy        214 -----------AGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGC 261
                          ||||||   ||||||  ||||||      ||||||
Db         52 SerSerSerIleProSerSerSerSerSerSerSerSerProSerSerSerHisSerSer 71

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAG 321
              ||||||   |||||||||   ||||||   :::         |||:::   :::
Db         72 SerSerProSerSerSerHisSerSerSerSerProSerSerSerSerSerThrSerSer 91

Qy        322 CCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGT 381
              |||   :::      |||      |||   ||||||   |||   |||
Db         92 ProSerSerSerSerSerSerSerSerSerSerSerProSerSerSerAsnSerSerSer 111

Qy        382 GCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCAC 441
              :::|||   ||||||         ||||||||||||   |||   ||||||   |||
Db        112 SerSerSerSerSerSerProSerSerSerSerSerSerSerSerSerSerProSerSer 131

Qy        442 TCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGC 501
              |||   ||||||   |||   ||||||||||||||||||         ||||||||||||
Db        132 SerSerSerSerProSerSerSerSerSerSerSerSerSerSerProSerSerSerSer 151

Qy        502 TTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTA 561
                      :::   |||:::      ::::::   :::|||
Db        152 SerSerProSerSerSerSerSerSerSerSerSerSerSerSerSerProSerSerSer 171

Qy        562 AACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGT 621
              :::||||||      |||   |||   |||||||||   :::|||      |||||||||
Db        172 SerProSerSerSerGlySerSerProSerSerSerAsnSerSerProSerSerSerSer 191

Qy        622 GGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGC 681
                    ||||||:::      |||                    ||||||
Db        192 SerSerProSerSerSerSerSerSerPro--------------------SerProArg 204

Qy        682 AGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCA 741
              |||      |||      |||   |||            ||||||:::   |||   |||
Db        205 SerSerSerProSerSerSerSerSerSerThrSerSerProSerThrSerSerProSer 224

Qy        742 GGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTG 801
                                                          |||   :::|||
Db        225 Ser--------------------------SerSerProSerSerSerSerPro------ 233

Qy        802 GTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACC 861
                       |||   :::      |||      :::|||      |||   :::
Db        234 ------SerSerSerCysProSerAlaAlaLeuGlyArgArgProGlnSerProGlnSer 251

Qy        862 TCT 864
              |||
Db        252 Ser 252
```

RESULT 2

US-11-096-568A-28959
; Sequence 28959, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28959
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1274)
; OTHER INFORMATION: Ceres Seq. ID no. 3095668
US-11-096-568A-28959

Alignment Scores:

Pred. No.:	1.69e-08	Length:	1274
Score:	262.50	Matches:	138
Percent Similarity:	36.5%	Conservative:	57
Best Local Similarity:	25.8%	Mismatches:	226
Query Match:	8.5%	Indels:	114
DB:	7	Gaps:	24

US-10-677-865A-2 (1-1590) x US-11-096-568A-28959 (1-1274)

```
Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTA---TCACCTCCCCTAACGACCCCTGCC 166
              |||::::::   ||||||   ||||||   :::   :::||||||:::      |||
Db        659 ProAlaLysProProValAlaThrProProIleAlaThrProProIleAlaLysPro--- 677

Qy        167 TCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCA 226
                                  |||   |||:::   |||       |||   :::
Db        678 -----------------ProValAlaThrProProThrAlaThrProProIleAlaThr 691

Qy        227 GCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCA 286
              :::|||         |||         |||      |||   :::            |||
Db        692 SerProValAlaLysProPro------ValAlaIleProProIleLysThrProProPro 709

Qy        287 GCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCA 346
              |||   |||   |||   :::      :::   :::                |||
Db        710 AlaLysProProValAlaIleProProIleAlaThrProProValAlaLysProProVal 729

Qy        347 GCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGA 406
              ||||||||||||                    |||:::
Db        730 AlaThrProProThrAlaThrProProIleAlaThrSerProIle-------------- 744

Qy        407 GCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGA 466

Db        744 ------------------------------------------------------------ 744

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTC 526
              |||                 :::         |||   |||                 :::
Db        745 AlaThrProProValValThrProProThrAlaThrSerProValAlaThrProProIle 764

Qy        527 TGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCT 586
                ::::|||   ||||||                  |||   |||:::
Db        765 AlaLysProProThrThr------------------ThrProProThrAlaThrProPro 778

Qy        587 CTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAA----------------- 628
              :::   :::|||   |||:::      |||      ||||||
Db        779 ValAlaMetPro---ProIleAlaThrProProThrAlaLysProProValAlaThrPro 797

Qy        629 ------GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCG------------ 670
                    |||:::   |||         |||   |||   |||   |||
Db        798 ProIleAlaAsnProProValGluLysProProValAlaThrProProIleAlaLysPro 817

Qy        671 ------ACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCA 724
                   ||||||   :::      |||         |||   :::||||||:::
```

```
Db        818 ProThrValLeuProProIleAlaLysProProValGluThr---SerProThrAlaThr 836

Qy        725 GCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTT 784
                ||||||                                          |||      |||
Db        837 ProProThrAla-----------------------------------ThrProProVal 844

Qy        785 CTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGC 844
                :::|||                 ||||||||||||:::
Db        845 AlaIleProPro------ValValLysProProValAlaIleProProIleThrLysPro 862

Qy        845 CCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCT 904
                ||||||       |||:::   |||||||||      |||          |||
Db        863 ProValAlaThrProProVal---ThrAsnProProThrAlaMetProProIleValThr 881

Qy        905 CCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGG 964
                |||      ::::::              |||   :::            |||
Db        882 ProProProIleValThr-----------ProProIleAlaLysSerProIle------ 895

Qy        965 GCTACTTCACCAAA-----------------GAGAACCCTGTGAAAGGCTCTCCAGGG 1005
                ||||||  |||                    :::   |||:::         |||
Db        896 AlaThrProProValSerThrProProIleAlaLysProProIleAlaThrProProVal 915

Qy       1006 GTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCC 1065
                  :::     |||   |||||||||:::                 ::::::||||||
Db        916 AlaThrThrProIleAlaLysProProIleAlaThrProPro------ThrAlaAsnPro 933

Qy       1066 ATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGT 1125
                :::   ::          |||   |||   ||||||      |||:::   |||
Db        934 ProValAlaAsnProProIleAlaLysSerProIleAlaLysProProIleAlaThr-Pr 953

Qy       1126 G-----------------------------------GGGTCCAGCTCTGTGGAGGCGGC 1149
                                                   |||      |||
Db        953 oProThrAlaMetProSerIleAlaThrProProIleGlyLysProProValAlaThrPr 973

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
                |||   |||   :::   |||          |||   :::   :::|||   ::::::
Db        973 oProMetAla---LysProProValAlaSerProProIleAlaThrProProIleIleLy 992

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
                  |||   |||            |||   |||            |||
Db        992 sProProValAlaThrProProIleThrLysProProValAlaThrProProValAlaTh 1012

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAAT---CGAGTGGCAAA 1326
                 |||     |||     |||   |||   :::|||:::              ||||||||
Db       1012 rProProIleAlaLysProProValAlaThrSerProIleGluThrProProValAlaLy 1032

Qy       1327 ATCATCCTTCAGCCTT------GTGGCAGCAAGTCCAGCTCTT---------CTGGTCAC 1371
                |         :::         ||||||      |||   :::          :::|||||
Db       1032 sProProValThrThrProProValAlaThrProProIleValLysProProIleValTh 1052

Qy       1372 CCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCT---CTCCCCAT 1428
                |            :::||||||               :::|||:::|||      |||||
Db       1052 rProProIleAlaThrProPro------IleAlaLysSerProIleAlaProProProIl 1070

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
                |     ||||||:::       |||||||||                      |||
Db       1070 eGlyThrProProIleAlaLysProProValAlaThrProProThrAlaThrProProVa 1090

Qy       1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCT 1548
                 |||   |||   |||:::    |||       |||:::      ||||||:::   |||
Db       1090 lAlaThrSerProIleAla-LysProProValAlaThrProProProAlaThrProProV 1110

Qy       1549 GAAGTTTTCCT---------ACCCCAAGGAGAGTTACT 1577
                  :::   |||         ||||||      ||||||
Db       1110 alAlaLysProProValAlaThrProProThrValThr 1122

RESULT 3
US-11-096-568A-28958
; Sequence 28958, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28958
;   LENGTH: 1480
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1480)
;   OTHER INFORMATION: Ceres Seq. ID no. 3095666
US-11-096-568A-28958

Alignment Scores:
Pred. No.:                1.67e-08        Length:          1480
Score:                    262.50          Matches:         138
Percent Similarity:       36.5%           Conservative:    57
Best Local Similarity:    25.8%           Mismatches:      226
Query Match:              8.5%            Indels:          114
DB:                       7               Gaps:            24

US-10-677-865A-2 (1-1590) x US-11-096-568A-28958 (1-1480)

Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTA---TCACCTCCCCTAACGACCCCTGCC 166
                |||::::::   ||||||  ||||||  :::   :::||||||:::      |||
Db        865 ProAlaLysProProValAlaThrProProIleAlaThrProProIleAlaLysPro--- 883

Qy        167 TCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCA 226
                                  |||   |||:::      |||      |||   :::
Db        884 ------------------ProValAlaThrProProThrAlaThrProProIleAlaThr 897

Qy        227 GCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCA 286
                :::|||          |||         |||      |||   :::          |||
Db        898 SerProValAlaLysProPro------ValAlaIleProProIleLysThrProProPro 915

Qy        287 GCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCA 346
                |||   |||   |||   :::      :::   :::                |||
Db        916 AlaLysProProValAlaIleProProIleAlaThrProProValAlaLysProProVal 935

Qy        347 GCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGA 406
                ||||||||||||                                ||||:::
Db        936 AlaThrProProThrAlaThrProProIleAlaThrSerProIle--------------- 950

Qy        407 GCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGA 466

Db        950 ------------------------------------------------------------ 950

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTC 526
                |||                  :::         |||   |||                :::
Db        951 AlaThrProProValValThrProProThrAlaThrSerProValAlaThrProProIle 970

Qy        527 TGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCT 586
                  :::|||   ||||||                     |||   |||:::
Db        971 AlaLysProProThrThr-----------------ThrProProThrAlaThrProPro 984

Qy        587 CTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAA------------------ 628
                :::   :::|||   |||:::     |||      ||||||
Db        985 ValAlaMetPro---ProIleAlaThrProProThrAlaLysProProValAlaThrPro 1003

Qy        629 ------GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCG------------ 670
                      |||:::   |||         |||   |||   |||   |||
Db       1004 ProIleAlaAsnProProValGluLysProProValAlaThrProProIleAlaLysPro 1023

Qy        671 ------ACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCA 724
                     ||||||   :::      |||         |||   :::||||||:::
Db       1024 ProThrValLeuProProIleAlaLysProProValGluThr---SerProThrAlaThr 1042

Qy        725 GCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTT 784
                  ||||||                                          |||      |||
Db       1043 ProProThrAla-----------------------------------ThrProProVal 1050
```

```
Qy        785 CTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGC 844
                 :::|||                |||||||||||:::
Db       1051 AlaIleProPro------ValValLysProProValAlaIleProProIleThrLysPro 1068

Qy        845 CCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCT 904
               ||||||         |||:::  |||||||||      |||           |||
Db       1069 ProValAlaThrProProVal---ThrAsnProProThrAlaMetProProIleValThr 1087

Qy        905 CCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGG 964
               |||      ::::::              |||   :::          |||
Db       1088 ProProProIleValThr-----------ProProIleAlaLysSerProIle------ 1101

Qy        965 GCTACTTCACCAAA-----------------GAGAACCCTGTGAAAGGCTCTCCAGGG 1005
               ||||||   |||                 :::   |||:::        |||
Db       1102 AlaThrProProValSerThrProProIleAlaLysProProIleAlaThrProProVal 1121

Qy       1006 GTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCC 1065
                  :::     |||  ||||||||:::                 ::::::||||||
Db       1122 AlaThrThrProIleAlaLysProProIleAlaThrProPro------ThrAlaAsnPro 1139

Qy       1066 ATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGT 1125
                 :::   :::         |||   |||   ||||||      |||:::   |||
Db       1140 ProValAlaAsnProProIleAlaLysSerProIleAlaLysProProIleAlaThr-Pr 1159

Qy       1126 G-----------------------------------GGGTCCAGCTCTGTGGAGGCGGC 1149
                                                   |||         |||
Db       1159 oProThrAlaMetProSerIleAlaThrProProIleGlyLysProProValAlaThrPr 1179

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
                |||   |||   :::   |||         |||   :::   :::|||   ::::::
Db       1179 oProMetAla---LysProProValAlaSerProProIleAlaThrProProIleIleLy 1198

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
                |||   |||         |||   |||         |||
Db       1198 sProProValAlaThrProProIleThrLysProProValAlaThrProProValAlaTh 1218

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAAT---CGAGTGGCAAA 1326
                |||      |||      |||   |||   :::|||:::            ||||||||
Db       1218 rProProIleAlaLysProProValAlaThrSerProIleGluThrProProValAlaLy 1238

Qy       1327 ATCATCCTTCAGCCTT------GTGGCAGCAAGTCCAGCTCTT---------CTGGTCAC 1371
                |      :::         |||||     |||  :::         :::|||||
Db       1238 sProProValThrThrProProValAlaThrProProIleValLysProProIleValTh 1258

Qy       1372 CCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCT---CTCCCCAT 1428
                |          :::||||||              :::|||:::|||      |||||
Db       1258 rProProIleAlaThrProPro------IleAlaLysSerProIleAlaProProProIl 1276

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
                |      ||||||:::         |||||||||                    |||
Db       1276 eGlyThrProProIleAlaLysProProValAlaThrProProThrAlaThrProProVa 1296

Qy       1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCT 1548
                |||   |||   |||:::    |||      |||:::      ||||||:::   |||
Db       1296 lAlaThrSerProIleAla-LysProProValAlaThrProProProAlaThrProProV 1316

Qy       1549 GAAGTTTTCCT---------ACCCCAAGGAGAGTTACT 1577
                  :::   |||        ||||||      ||||||
Db       1316 alAlaLysProProValAlaThrProProThrValThr 1328
```

RESULT 4
US-11-096-568A-28960
; Sequence 28960, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28960
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1267)
; OTHER INFORMATION: Ceres Seq. ID no. 3095669
US-11-096-568A-28960

Alignment Scores:

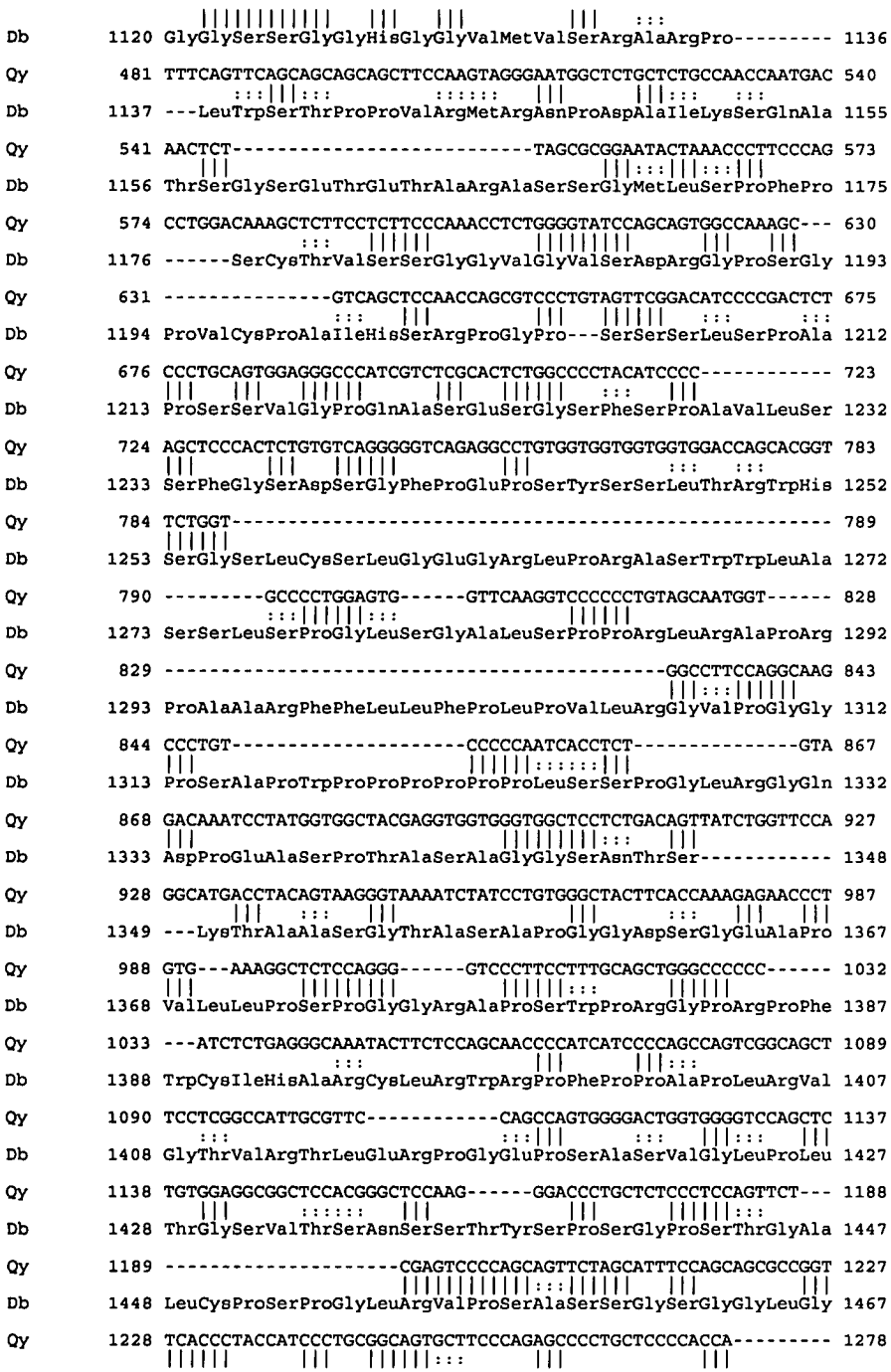
Pred. No.:	1.81e-08	Length:	1267
Score:	262.00	Matches:	138
Percent Similarity:	36.5%	Conservative:	57
Best Local Similarity:	25.8%	Mismatches:	226
Query Match:	8.4%	Indels:	114
DB:	7	Gaps:	24

US-10-677-865A-2 (1-1590) x US-11-096-568A-28960 (1-1267)

```
Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTA---TCACCTCCCCTAACGACCCCTGCC 166
                |||::::::   ||||||   ||||||   :::   :::||||||:::      |||
Db        652 ProAlaLysProProValAlaThrProProIleAlaThrProProIleAlaLysPro--- 670

Qy        167 TCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCA 226
                              |||  ||||:::   |||      |||   :::
Db        671 -----------------ProValAlaThrProProThrAlaThrProProIleAlaThr 684

Qy        227 GCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCA 286
                :::|||         |||         |||   :::                |||
Db        685 SerProValAlaLysProPro------ValAlaIleProProIleLysThrProProPro 702

Qy        287 GCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCA 346
                |||   |||   |||   :::       :::   :::                |||
Db        703 AlaLysProProValAlaIleProProIleAlaThrProProValAlaLysProProVal 722

Qy        347 GCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGA 406
                ||||||||||||                ||:::
Db        723 AlaThrProProThrAlaThrProProIleAlaThrSerProIle--------------- 737

Qy        407 GCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGA 466

Db        737 ------------------------------------------------------------ 737

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTC 526
                |||                 :::         |||   |||                 :::
Db        738 AlaThrProProValValThrProProThrAlaThrSerProValAlaThrProProIle 757

Qy        527 TGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCT 586
                  :::|||   ||||||                 |||   |||:::
Db        758 AlaLysProProThrThr-----------------ThrProProThrAlaThrProPro 771

Qy        587 CTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAA----------------- 628
                :::   :::|||   |||:::      |||      ||||||
Db        772 ValAlaMetPro---ProIleAlaThrProProThrAlaLysProProValAlaThrPro 790

Qy        629 ------GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCG------------ 670
                    |||:::   |||         |||   |||   |||   |||
Db        791 ProIleAlaAsnProProValGluLysProProValAlaThrProProIleAlaLysPro 810

Qy        671 ------ACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCA 724
                     ||||||   :::      |||         |||   :::||||||:::
Db        811 ProThrValLeuProProIleAlaLysProProValGluThr---SerProThrAlaThr 829

Qy        725 GCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTT 784
                  ||||||                                        |||      |||
Db        830 ProProThrAla-----------------------------------ThrProProVal 837

Qy        785 CTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGC 844
                  :::|||                |||||||||||:::
Db        838 AlaIleProPro------ValValLysProProValAlaIleProProIleThrLysPro 855

Qy        845 CCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCT 904
               ||||||         |||:::  |||||||||      |||           |||
```

```
Db        856 ProValAlaThrProProVal---ThrAsnProProThrAlaMetProProIleValThr 874

Qy        905 CCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGG 964
              |||      ::::::               |||   :::            |||
Db        875 ProProProIleValThr-----------ProProIleAlaLysSerProIle------ 888

Qy        965 GCTACTTCACCAAA-----------------GAGAACCCTGTGAAAGGCTCTCCAGGG 1005
              ||||||   |||                    :::   |||:::         |||
Db        889 AlaThrProProValSerThrProProIleAlaLysProProIleAlaThrProProVal 908

Qy       1006 GTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCC 1065
                 :::      |||   |||||||||:::              ::::::||||||
Db        909 AlaThrThrProIleAlaLysProProIleAlaThrProPro------ThrAlaAsnPro 926

Qy       1066 ATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGT 1125
                :::   :::         |||   |||   ||||||      |||:::   |||
Db        927 ProValAlaAsnProProIleAlaLysSerProIleAlaLysProProIleAlaThr-Pr 946

Qy       1126 G-----------------------------------GGGTCCAGCTCTGTGGAGGCGGC 1149
                                                   |||      |||
Db        946 oProThrAlaMetProSerIleAlaThrProProIleGlyLysProProValAlaThrPr 966

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
                |||   |||   :::   |||         |||   :::   :::|||   ::::::
Db        966 oProMetAla---LysProProValAlaSerProProIleAlaThrProProIleIleLy 985

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
                |||   |||             |||   |||             |||
Db        985 sProProValAlaThrProProIleThrLysProProValAlaThrProProValAlaTh 1005

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAAT---CGAGTGGCAAA 1326
                |||      |||      |||   |||   :::|||:::              |||||||
Db       1005 rProProIleAlaLysProProValAlaThrSerProIleGluThrProProValAlaLy 1025

Qy       1327 ATCATCCTTCAGCCTT------GTGGCAGCAAGTCCAGCTCTT---------CTGGTCAC 1371
                |        :::         ||||||      |||   :::         :::|||||
Db       1025 sProProValThrThrProProValAlaThrProProIleValLysProProIleValTh 1045

Qy       1372 CCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCT---CTCCCCAT 1428
                |            :::||||||               :::|||:::|||      |||||
Db       1045 rProProIleAlaThrProPro------IleAlaLysSerProIleAlaProProProIl 1063

Qy       1429 CCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
                |      ||||||:::          ||||||||                  |||
Db       1063 eGlyThrProProIleAlaLysProProValAlaThrProProThrAlaThrProProVa 1083

Qy       1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCT 1548
                |||   |||   |||:::   |||       |||:::      ||||||:::   |||
Db       1083 lAlaThrSerProIleAla-LysProProValAlaThrProProProAlaThrProProV 1103

Qy       1549 GAAGTTTTCCT---------ACCCCAAGGAGAGTTACT 1577
                  :::   |||         ||||||      ||||||
Db       1103 alAlaLysProProValAlaThrProProThrValThr 1115

RESULT 5
US-10-821-234-1070
; Sequence 1070, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1070
;   LENGTH: 1377
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1070

Alignment Scores:
Pred. No.:              3.37e-07        Length:         1377
Score:                  240.00          Matches:        177
Percent Similarity:     33.4%           Conservative:   55
Best Local Similarity:  25.5%           Mismatches:     228
Query Match:            7.7%            Indels:         235
DB:                     6               Gaps:           34

US-10-677-865A-2 (1-1590) x US-10-821-234-1070 (1377)

Qy          12 TCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGG---ATGATGGCACTGCTGCTGGC 68
               |||   ||||||   |||   ||| ||||||   |||      :::||||||||||   ||
Db        442 SerHisThrLeuSerGlyGlyCys-GlyGlyArgGlySerLysValAlaLeuLeuProAl 461

Qy          69 TGGTCTCCTCCTG---------------------------------------------- 81
               ||||     |||
Db        461 aGlyGlyAlaLeuGlnHisSerArgSerMetSerMetProValAlaHisSerProProAl 481

Qy          82 ----------CCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGA 131
                         ||||||:::|||:::   |||   |||      |||
Db        481 aAlaThrSerProGlySerLeuSerSerSerSerGlyHisGlySerGlySerTyrProPr 501

Qy         132 CCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC-----------------ACTGG 173
               |||                 |||:::   |||                      ||
Db        501 oPro---------ProGlyProHisProProLeuProHisProLeuHisHisGlyProGl 518

Qy         174 GAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGC----- 228
               |:::     |||||||||   :::|||   ||||||||   |||
Db        518 yGlnArgProSerSerGlySerAlaSerAlaSerGlySerProSerAspProGlyPheMe 538

Qy         229 -TCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGC--------------TCCAGTGGTAG 272
               |||:::              ||||||   |||                 |||      ||
Db        538 tSerLeuAspGluTyrGlySerSerProGlyAspLeuArgAlaPheCysSerHisArgSe 558

Qy         273 CTCCAGCGGATCCAGCATTGCCCAG-----------------GGTGGTTCTGCAGGATC 314
               |::::::      |||||||||:::                 ||||||      |||
Db        558 rAsnThrProGluSerIleAlaGluThrProProAlaArgAspGlyGlyGlyGlyGlyGl 578

Qy         315 TTTT--------------------AAGCCA------------GGAACGGGGTATTCCCA 341
               |||                    :::|||            |||      |||   ::
Db        578 uPheTyrGlyTyrMetThrMetAspArgProLeuSerHisCysGlyArgSerTyrArgAr 598

Qy         342 GGTCAGC-------------------------------------------------TACTCC-- 354
               :||||||                                                 ||||||
Db        598 gValSerGlyAspAlaAlaGlnAspLeuAspArgGlyLeuArgLysArgThrTyrSerLe 618

Qy         355 ---------------------------------TCCGGATCTGGCTCT----------- 369
                                                  |||   |||
Db        618 uThrThrProAlaArgGlnArgProValProGlnProSerSerAlaSerLeuAspGluTy 638

Qy         370 -ACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAG 428
                ||||||      |||:::   |||          |||                  ::
Db        638 rThrLeuMetArgAlaThrPheSerGlySerAlaGlyArgLeuCysProSerCysProAl 658

Qy         429 CAGC-------------------------------------------------------GG 434
               :|||                                                       ||
Db        658 aSerSerProLysValAlaTyrHisProTyrProGluAspTyrGlyAspIleGluIleGl 678

Qy         435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
               |||||||      |||||||||:::      :::                     :::
Db        678 ySerHisArgSerSerSerSerAsnLeuGlyAlaAspAspGlyTyrMetProMetThrPr 698

Qy         495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCT---------------------------- 525
                  :::      |||:::||||||
Db        698 oGlyAlaAlaPheAlaGlySerGlySerGlySerCysArgSerAspAspTyrMetProMe 718

Qy         526 -CTGCCAACCAATGACAACTCTTAGCGCGGAATACTA----------------------- 561
                   |||   :::   ::::::   :::   ||||||
Db        718 tSerProAlaSerValSerAlaProLysGlnIleLeuGlnProArgAlaAlaAlaAlaAl 738
```

```
Qy       562 ----------------------AACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 599
                                   |||:::   |||      :::|||:::::   ::
Db       738 aAlaAlaAlaValProSerAlaGlyProAlaGlyProAlaProThrSerAlaAlaGlyAr 758

Qy       600 AACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTC 659
             :|||         |||      |||         ||||||         |||      ||
Db       758 gThrPheProAlaSerGlyGlyGlyTyrLysAlaSerSer---------ProAlaGluSe 775

Qy       660 GGACATCCCCGACTCTCCC--------------TGCAGTGGAGGGCCCATCGTCTCGCA 704
             |         ||||||                |||                :::   ||
Db       775 rSerProGluAspSerGlyTyrMetArgMetTrpCysGlySerLysLeuSerMetGluHi 795

Qy       705 CTCT---GGCCCCTACATCCCCAGC-----------------------TCCCACTCTGT 737
             |:::   |||     :::|||:::                       |||   :::||
Db       795 sAlaAspGlyLysLeuLeuProAsnGlyAspTyrLeuAsnValSerProSerAspAlaVa 815

Qy       738 GTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCT-- 795
             |:::   |||      |||                        |||      |||   |||
Db       815 lThrThrGlyThrProProAspPhePheSerAlaAlaLeuHisProGlyGlyGluProLe 835

Qy       796 ----GGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCC 851
                 ||||||         |||   |||            ||||||
Db       835 uArgGlyVal---------ProGlyCysCysTyrSerSerLeuPro-------------- 847

Qy       852 CCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGA 911
                           :::||||||                    ||||||   |||||
Db       848 ------------------ArgSerTyrLysAlaProTyrThrCysGlyGlyAspSerAs 861

Qy       912 CAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTT 971
             |   |||::::::                                 |||||||||
Db       861 pGlnTyrValLeuMetSerSer-----------------------ProValGlyArgIl 873

Qy       972 CACCAAAGAGAACCCTGTG------AAAGGCTCTCCAGGG-----------GTCCCTTC 1013
                 :::|||      :::      :::   :::||||||                ::
Db       873 eLeuGluGluGluArgLeuGluProGlnAlaThrProGlyProSerGlnAlaAlaSerAl 893

Qy       1014 CTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCC 1073
              :|||   |||||||||                    :::|||::::::||
Db       893 aPheGlyAlaGlyProThr-----------------GlnProProHisProValValPr 907

Qy       1074 CAGC-----------------------------------------CAGTCGGCAGCTTC 1091
              ||||                                       :::      |||
Db       907 oSerProValArgArgSerGlyArgArgSerGlyPheLeuGlyGlnArgGlyArgAla-V 927

Qy       1092 CTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTC 1151
                ||||||   |||   |||      ||||||              |||:::          |
Db       927 alArgProThrArgLeuSerLeuGluGlyLeuProSerLeuProSerMetHisGluTyrP 947

Qy       1152 CACGGGCTCCAAGG---------------------------------------------- 1165
              ||      |||
Db       947 roLeuProProGluProLysSerProGlyGluTyrIleAsnIleAspPheGlyGluProG 967

Qy       1166 -----------GACCCTGCTCTCCCTCCAGTTCTCGAGTCCCC-AGCAGTTCTAGCATTT 1213
                        |||      ||||||:::|||   |||   :::|||||||||
Db       967 lyAlaArgLeuSerProProAlaProProLeuLeuAlaSerAlaAlaSerSerSerSerL 987

Qy       1214 CCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCC 1273
              ||||||   ||||||                ||||||   :::
Db       987 euLeuSerAlaSerSerProAlaSerSerLeuGlySerGlyThr---------------- 1001

Qy       1274 CACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCC 1333
                |||||||||   |||      |||      |||      |||      :::    |
Db       1002 --ProGlyThrSerSer---AspSerArgGlnArgSerProLeuSerAspTyrMetAsnL 1020

Qy       1334 TTCAGCCTTGTGGCAGCAAGTCC-----------AGCTCTTCTGGTCACCCTTGCATGT 1381
              ||                |||||           :::   ||||||||||||      |
Db       1020 euAspPheSerSerProLysSerProLysProGlyAlaProSerGlyHisProValGlyS 1040

Qy       1382 CTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCG 1441
              ||:::      |||                ||||:::   |||      |||    ||||||
Db       1040 erLeuAspGlyLeuLeu---------SerProGluAlaSer---SerPro-TyrProPro 1055

Qy       1442 CTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGG 1501
              |||   |||   |||   |||                          |||      ||| |||
Db       1056 LeuProProArgProSerAlaSerProSerSerSerLeuGlnProProProPro-Pro-- 1074

Qy       1502 ATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCA 1535
                 |||:::|||   |||           |||
Db       1075 ----ProAlaProGlyGluLeuTyrArgLeuPro 1084

RESULT 6
US-11-182-016-23
; Sequence 23, Application US/11182016
; Publication No. US20060019294A1
; GENERAL INFORMATION:
;  APPLICANT: SUGEN, INC.
;  TITLE OF INVENTION: TYROSINE KINASE SUBSTRATE (TKS) PROTEINS
;  FILE REFERENCE: 038602/0102
;  CURRENT APPLICATION NUMBER: US/11/182,016
;  CURRENT FILING DATE:  2005-07-15
;  PRIOR APPLICATION NUMBER: US/09/958,359
;  PRIOR FILING DATE: 2002-02-05
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 1742
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Grub
US-11-182-016-23

Alignment Scores:
Pred. No.:              3.52e-07        Length:         1742
Score:                  239.50          Matches:        151
Percent Similarity:     34.8%           Conservative:   68
Best Local Similarity:  24.0%           Mismatches:     246
Query Match:            7.7%            Indels:         165
DB:                     7               Gaps:           28

US-10-677-865A-2 (1-1590) x US-11-182-016-23 (1-1742)

Qy         46 GGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATT 105
              |||      :::         :::||||||   :::|||   |||   :::   |||
Db        965 GlySerAlaSerThrPheCysSerGlyLeuGlyIlePro---ThrAlaSerSerSerSer 983

Qy        106 GGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCC--- 162
              |||:::   :::                        |||   |||:::   |||
Db        984 GlySerAlaAlaProSerSerSerSerMetAlaCysArgThrGluProSerThrProTrp 1003

Qy        163 ------------------------------TGCCTCACT---------GGGAAGGGT--- 180
                                           |||::::::         |||:::|||
Db       1004 HisArgArgHisSerLeuThrThrSerTrpCysIleSerThrGlnLeuGlyGluGlySer 1023

Qy        181 ---------------------GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
                                  |||||||||   ::::::          |||   |||
Db       1024 LeuProProAsnSerHisCysSerSerSerGlyLeuAlaThrLeuPheArgSerPheSer 1043

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT------------ 267
                 :::|||      :::   :::||||||   |||   ||||||
Db       1044 AspSerThrSerAlaProAsnProGlnSerServalGlyArgSerSerMetArgIleSer 1063

Qy        268 ---------------GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
                              ||||||          |||   :::      |||   |||||||||
Db       1064 SerArgThrArgGlyGlySerProGluSerSerTrpLeuLysAlaGlyMetSerAlaGly 1083

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372
                  |||:::|||   ::::::   |||       :::   |||   |||
Db       1084 ------------GlySerGlyGlyAlaArgArgSerSerArgThrAspSerProSerVal 1099

Qy        373 CTACAAGGTGCATCCGGTTCCTCC------------CAGCTGGGGAGCAGCAGCTCTCAC 420
                 :::           |||                          |||
Db       1100 CysArgSerGlyLeuSerSerGluTrpArgArgValValPheSerValSerArgGluGly 1119

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
```

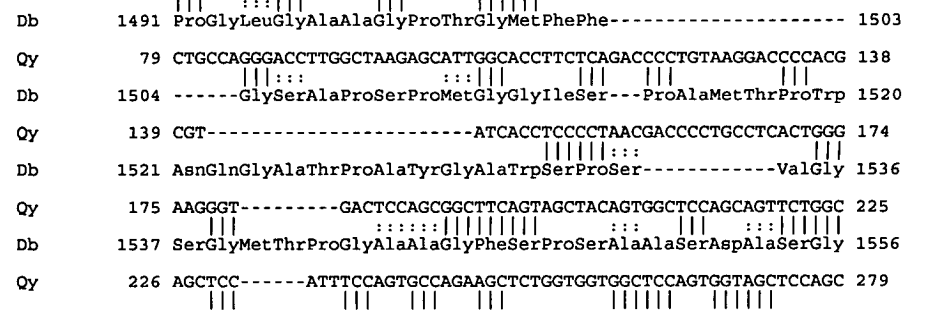

Db 1120 GlyGlySerSerGlyGlyHisGlyGlyValMetValSerArgAlaArgPro--------- 1136
Qy 481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db 1137 ---LeuTrpSerThrProProValArgMetArgAsnProAspAlaIleLysSerGlnAla 1155
Qy 541 AACTCT--------------------------TAGCGCGGAATACTAAACCCTTCCCAG 573
Db 1156 ThrSerGlySerGluThrGluThrAlaArgAlaSerSerGlyMetLeuSerProPhePro 1175
Qy 574 CCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGC--- 630
Db 1176 ------SerCysThrValSerSerGlyGlyValGlyValSerAspArgGlyProSerGly 1193
Qy 631 --------------GTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCT 675
Db 1194 ProValCysProAlaIleHisSerArgProGlyPro---SerSerSerLeuSerProAla 1212
Qy 676 CCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC------------ 723
Db 1213 ProSerSerValGlyProGlnAlaSerGluSerGlySerPheSerProAlaValLeuSer 1232
Qy 724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
Db 1233 SerPheGlySerAspSerGlyPheProGluProSerTyrSerSerLeuThrArgTrpHis 1252
Qy 784 TCTGGT-- 789
Db 1253 SerGlySerLeuCysSerLeuGlyGluGlyArgLeuProArgAlaSerTrpTrpLeuAla 1272
Qy 790 ---------GCCCCTGGAGTG------GTTCAAGGTCCCCCCTGTAGCAATGGT------ 828
Db 1273 SerSerLeuSerProGlyLeuSerGlyAlaLeuSerProProArgLeuArgAlaProArg 1292
Qy 829 ---GGCCTTCCAGGCAAG 843
Db 1293 ProAlaAlaArgPhePheLeuLeuPheProLeuProValLeuArgGlyValProGlyGly 1312
Qy 844 CCCTGT--------------------CCCCCAATCACCTCT--------------GTA 867
Db 1313 ProSerAlaProTrpProProProProProProLeuSerSerProGlyLeuArgGlyGln 1332
Qy 868 GACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCA 927
Db 1333 AspProGluAlaSerProThrAlaSerAlaGlyGlySerAsnThrSer------------ 1348
Qy 928 GGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCT 987
Db 1349 ---LysThrAlaAlaSerGlyThrAlaSerAlaProGlyGlyAspSerGlyGluAlaPro 1367
Qy 988 GTG---AAAGGCTCTCCAGGG------GTCCCTTCCTTTGCAGCTGGGCCCCCC------ 1032
Db 1368 ValLeuLeuProSerProGlyGlyArgAlaProSerTrpProArgGlyProArgProPhe 1387
Qy 1033 ---ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
Db 1388 TrpCysIleHisAlaArgCysLeuArgTrpArgProPheProProAlaProLeuArgVal 1407
Qy 1090 TCCTCGGCCATTGCGTTC------------CAGCCAGTGGGGACTGGTGGGGTCCAGCTC 1137
Db 1408 GlyThrValArgThrLeuGluArgProGlyGluProSerAlaSerValGlyLeuProLeu 1427
Qy 1138 TGTGGAGGCGGCTCCACGGGCTCCAAG------GGACCCTGCTCTCCCTCCAGTTCT--- 1188
Db 1428 ThrGlySerValThrSerAsnSerSerThrTyrSerProSerGlyProSerThrGlyAla 1447
Qy 1189 --------------------CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGT 1227
Db 1448 LeuCysProSerProGlyLeuArgValProSerAlaSerSerGlySerGlyGlyLeuGly 1467
Qy 1228 TCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCA--------- 1278
Db 1468 SerProGlyGlyProGluGlySerSerGlyGlySerValLeuGlyCysProThrMetArg 1487
Qy 1279 --------------GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCC------------ 1311
Db 1488 ProSerLeuCysThrAsnArgGlyGlnMetSerSerAlaGlnGlySerArgIleProGlu 1507
Qy 1312 --------------------------CAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
Db 1508 GlyAlaAspSerArgGlnValLeuArgArgProValGlyArgSerLeuLeuMetProCys 1527
Qy 1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
Db 1528 SerHisSerGlyValAsnArgAlaGlnAlaSerLeuGlyThrGlyThr---SerCysThr 1546
Qy 1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
Db 1547 ArgProProProGlyAlaArgHisLeuSerAlaSerArgGlyAlaTrpAlaAlaGluGly 1566
Qy 1465 AGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494
Db 1567 ThrThrCysArgLysSerProGlyArgSer 1576

RESULT 7
US-10-821-234-1641
; Sequence 1641, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1641
; LENGTH: 1970
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1641

Alignment Scores:

Pred. No.:	6.77e-07	Length:	1970
Score:	234.50	Matches:	154
Percent Similarity:	39.3%	Conservative:	58
Best Local Similarity:	28.6%	Mismatches:	212
Query Match:	7.6%	Indels:	115
DB:	6	Gaps:	26

US-10-677-865A-2 (1-1590) x US-10-821-234-1641 (1-1970)


Qy 19 CCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTC 78
Db 1491 ProGlyLeuGlyAlaAlaGlyProThrGlyMetPhePhe---------- 1503
Qy 79 CTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACG 138
Db 1504 ------GlySerAlaProSerProMetGlyGlyIleSer---ProAlaMetThrProTrp 1520
Qy 139 CGT----------------------ATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
Db 1521 AsnGlnGlyAlaThrProAlaTyrGlyAlaTrpSerProSer----------ValGly 1536
Qy 175 AAGGGT---------GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGC 225
Db 1537 SerGlyMetThrProGlyAlaAlaGlyPheSerProSerAlaAlaSerAspAlaSerGly 1556
Qy 226 AGCTCC------ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279

```
Db       1557 PheSerProGlyTyrSerProAlaTrpSerProThrProGlySerProGlySerProGly 1576

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              ||||||                  |||:::      |||           :::
Db       1577 ProSerSerProTyrIleProSerProGlyGlyAlaMetSerProSerTyrSerProThr 1596

Qy        340 CAGGTCAGCTACTCC------TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCC 393
                :::|||        |||      |||  |||  |||      |||  |||
Db       1597 SerProAlaTyrGluProArgSerProGlyGlyTyrThrProGlnSerProSerTyrSer 1616

Qy        394 TCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGC 453
                    :::|||   |||:::|||   :::|||   |||:::|||   :::|||
Db       1617 Pro---------ThrSerProSerTyrSerProThrSerProSerTyrSerProThrSer 1633

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGG 513
                 ::::::|||   :::|||   |||:::      :::|||   |||:::
Db       1634 ProAsnTyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThr 1653

Qy        514 AATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAG 573
               :::   |||      ||||||:::   :::                 :::|||:::
Db       1654 SerProSerTyrSerProThrSerProSer--------------TyrSerProThrSer 1668

Qy        574 CCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTC 633
              |||      |||   :::|||        |||      |||   |||
Db       1669 ProSerTyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThr 1688

Qy        634 AGCTCCAACCAGCGTCCCTGTAGT---TCGGACATCCCCGACTCTCCCTGCAGTGGAGGG 690
              |||   :::      |||   |||   |||    |||||||||   |||:::
Db       1689 SerProSerTyrSerProThrSerProSer-TyrSerProThrSerProSerTyr---Se 1707

Qy        691 CCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAG 750
               |||:::|||   :::   :::||||||||||||:::
Db       1707 rProThrSerProSerTyrSerProThrSerProSer---------------------- 1719

Qy        751 AGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGT 810
                        :::      ||||||       :::   |||   |||   |||   |
Db       1720 ------------TyrSerProThrSerPro-SerTyrSerProThrSerProSerTyrS 1735

Qy        811 CCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT---GTA 867
              |||||                |||||||           |||||||||::: |||
Db       1735 erProThr---------SerProSerTyrSerProThrSerProAsnTyr-SerProThr 1751

Qy        868 GACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCA 927
                 :::|||                              :::|||   ||||||
Db       1752 SerProAsnTyr-----------------ThrProThrSerProSerTyrSerProThr 1765

Qy        928 GGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGC---TACTTCACCAAAGAGAAC 984
                 :::||||||                |||           :::|||      :::
Db       1766 SerProSerTyrSer--------------ProThrSerProAsnTyrThrProThrSer 1780

Qy        985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
              |||         ||||||      ||||||::::::         |||
Db       1781 Pro---AsnTyrSerProThrSerProSerTyrSerProThrSerPro---------- 1795

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
                    |||   :::|||      ||||||   |||   :::|||
Db       1796 ---------SerTyrSerProThrSerProSerTyrSerProSerSerProArgTyrThr 1812

Qy       1105 TTCCAG---CCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCC 1161
                 |||   |||      |||                    |||:::   |||
Db       1813 ProGlnSerProThrTyrThr-----------------------ProSerSerProSer 1824

Qy       1162 AAGGGACCCTGCTCTCCCTCCAGTTCT--------------CGAGTCCCCAGCAGTTCT 1206
                 |||   |||||||||   |||                  |||:::|||
Db       1825 TyrSerProSerSerProSerTyrSerProThrSerProLysTyrThrProThrSerPro 1844

Qy       1207 AGCATTTCCAGCAGCGCC----------------------------------GGTTCA 1230
              |||   |||   |||:::                                   |||
Db       1845 SerTyrSerProSerSerProGluTyrThrProThrSerProLysTyrSerProThrSer 1864

Qy       1231 CCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCA---------CCAGGC 1281
              |||   :::      |||      ||||||   ||||||         |||
Db       1865 ProLysTyrSerProThrSerProLysTyrSerProThrSerProThrTyrSerProThr 1884

Qy       1282 ACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCT 1341
                 |||      :::|||   :::|||   :::   |||            :::      |||
Db       1885 ThrProLysTyrSerProThrSerProThrTyrSerProThrSerProValTyrThrPro 1904

Qy       1342 TGTGGCAGCAAGTCCAGC-----------------TCTTCTGGTCACCCTTGCATGTCT 1383
                    |||   |||                  |||           |||      |||
Db       1905 ThrSerProLysTyrSerProThrSerProThrTyrSerProThrSerProLysTyrSer 1924

Qy       1384 GTCTCCTCCTTGACACTGACT---GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
                 :::|||   |||   :::          |||   ||||||   :::          |||
Db       1925 ProThrSerProThrTyrSerProThrSerProLysGlySerThrTyrSerProThrSer 1944

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCC 1489
                 |||    ||||||   :::|||                |||||||||
Db       1945 ProGlyTyr-SerProThrSerProThrTyrSerLeuThrSerProAla 1960

RESULT 8
US-11-087-099-7521
; Sequence 7521, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7521
;   LENGTH: 7465
;   TYPE: PRT
;   ORGANISM: Magnetococcus sp. MC-1
US-11-087-099-7521

Alignment Scores:
Pred. No.:              8.34e-07      Length:          7465
Score:                  232.00        Matches:         115
Percent Similarity:     27.1%         Conservative:    13
Best Local Similarity:  24.4%         Mismatches:      162
Query Match:            7.5%          Indels:          182
DB:                     7             Gaps:            15

US-10-677-865A-2 (1-1590) x US-11-087-099-7521 (1-7465)

Qy       1497 GATGGAGCGGCAGGGGATCTTTCCAGCACTGCTGGAGCCACAGGGCTTGGCACCAGCGGA 1438
              |||||||||   |||||||||                |||||      |||  :::|||
Db       6920 AspGlyAlaProGlyAspLeuLeuGlyAlaGlyGlyAlaAlaAlaGlyGlyAlaThrGly 6939

Qy       1437 GGGATCAGGATGGGGAGAGCCATCGGGGCCCCCAGTCAGTGTCAAGGAGGAGACAGACAT 1378
              |||          |||          |||      :::          ||||||
Db       6940 GlyThrThrGlyGly---------GlyThrThrAlaGlyThrGlyGlyGly--------- 6953

Qy       1377 GCAAGGGTGACCAGAAGAGCTGGACTTGCTGCCACAAGGCTGAAGGATGATTTTGCCACT 1318
                                      ||||||
Db       6954 ---------------------GlyLeuGlyGlyGly------------------------ 6958

Qy       1317 CGATTGGGAACTGGAGCTGCTGCTGAAGGAGCCGGTGCCTGGTGGGGAGCAGGGGCTCTG 1258
                    |||   |||:::   |||           |||           |||
Db       6959 ------GlyPheGlySerGlyAlaProAsnThrSerAlaGlyGlyLeuThrGlyGlyPhe 6976

Qy       1257 GGAAGCACTGCCGCAGGG------ATGGTAGGGTGAACCGGCGCTGCTGGAAATGCTAGA 1204
              |||   ||||||   |||      :::   |||   ||||||      |||   |||
Db       6977 GlyGlyThrAlaThrGlyGlyThrValThrGlyGlyThrGlyGlyThrGlyGlyAlaAla 6996

Qy       1203 ACTGCTGGGGACTCGAGAACTGGAGGGAGAGCAGGGTCCCTTGGAGCCCGTGGAGCCGCC 1144
              |||   |||   :::      |||          |||      |||      |||
Db       6997 ThrGlyGlyThrAlaThrGlyGlyThrAlaThrGlyGlyThrGlyGlyThrGlyGlyThr 7016

Qy       1143 TCCACAGAGCTGGACCCCACCAGTCCCCACTGGCTGGAACGCAATGGCCGAGGAAGCTGC 1084
              :::|||          |||                           |||
Db       7017 ThrThrGlyGlyThrAlaThr---------------------------GlyGlyThr 7026

Qy       1083 CGACTGGCTGGGGATGATGGGGTTGCTGGAGAAGTATTTGCCCTCAGAGATGGGGGGCCC 1024
```

```
Db       7027 AlaThrGlyGlyThrGlyGlyAlaAlaGlyGlyThrGlyGlyAlaAlaAlaGlyGlyAla 7046

Qy       1023 AGCTGCAAAGGAAGGGACCCCTGGAGAGCCTTTCACAGGGTTCTCTTTGGTGAAGTAGCC 964

Db       7047 AlaThr----------------------------------------GlyGluAlaGly 7052

Qy        963 CACAGGATAGATTTTACCCTTACTGTAGGTCATGCCTGGAACCAGATAACTGTCAGAGGA 904

Db       7052 ------------------------------------------------------------ 7052

Qy        903 GCCACCCACCACCTCGTAGCCACCATAGGATTTGTCTACAGAGGTGATTGGGGGACAGGG 844

Db       7053 -------------------------GlyThrAlaThrGlyGlyThrGlyGlyThrGly 7063

Qy        843 CTTGCCTGGAAGGCCACCATTGCTACAGGGGGGACCTTGAACCACTCCAGGGGCACCAGA 784

Db       7064 GlyAlaAlaThrGlyGlyThrAlaThrGlyGlyThrGlyGlyAlaAlaGlyGlyThrGly 7083

Qy        783 ACCGTGCTGGTCCACCACCACCACCACAGGCCTCTGACCCCCTGACACAGAGTGGGAGCT 724

Db       7084 ------------------------------------------------------GlyAla 7085

Qy        723 GGGGATGTAGGGGCCAGAGTGCGAGACGATGGGCCCTCCACTGCAGGGAGAGTCGGGGAT 664

Db       7086 AlaAlaGlyGlyAlaAlaThrGlyGluAlaGlyGlyThrAlaThrGlyGlyThrGlyGly 7105

Qy        663 GTCCGAACTACAGGGACGCTGGTTGGAGCTGACGCTTTGGCCACTGCTGGATACCCCAGA 604

Db       7106 ThrAlaThrGlyGlyThrAlaThrGly------------GlyThrGlyGlyAlaAlaGly 7121

Qy        603 GGTTTGGGAAGAGGAAGAGCTTTGTCCAGGCTGGGAAGGGTTTAGTATTCCGCGCTAAGA 544

Db       7122 GlyThrGly---------------------------------------------------- 7124

Qy        543 GTTGTCATTGGTTGGCAGAGCAGAGCCATTCCCTACTTGGAAGCTGCTGCTGCTGAACTG 484

Db       7124 ------------------------------------------------------------ 7124

Qy        483 AAAGCTGCTGCTGCTGCTCGAATGAGAGCTGCTGCTTCCCGAGTGAGAGCCGCTGCTTCC 424

Db       7125 ---------GlyAlaAlaAlaGlyGlyAlaAlaThrGlyGluAlaGlyGlyAlaAlaThr 7141

Qy        423 CGAGTGAGAGCTGCTGCTCCCCAGCTGGGAGGAACCGGATGCACCTTGTAGAGTAGAGCC 364

Db       7142 -------------------------GlyGlyThrGlyGlyThr--------------- 7147

Qy        363 AGATCCGGAGGAGTAGCTGACCTGGGAATACCCCGTTCCTGGCTTAAAAGATCCTGCAGA 304

Db       7148 AlaThrGlyGlyThrAla---------------------------------------- 7153

Qy        303 ACCACCCTGGGCAATGCTGGATCCGCTGGAGCTACCACTGGAGCCACC------------ 256

Db       7154 ---ThrGlyGlyThrGlyGlyAlaAlaGlyGlyThrGlyGlyAlaAlaAlaGlyGlyAla 7172

Qy        255 --------------------ACCAGAGCTTCTGGCACTGGAAATGGAGCTGCCAGAACT 217

Db       7173 AlaThrGlyGluAlaGlyGlyAlaAlaThrGlyGlyThrGlyGlyAlaAlaAlaGlyGly 7192

Qy        216 GCTGGAGCCACTGTAGCTACTGAAGCCGCTGGAGTCACCCTTCCCAGTGAGGCAGGGGTC 157

Db       7193 ThrGlyGlyAlaAlaThrGlyGluAlaGlyGlyAlaAlaThrGlyGlyThrAlaGly--- 7211

Qy        156 GTTAGGGGAGGTGATACGCGTGGGGTCCTTACAGGG 121

Db       7212 ---AspAlaGlyAspThrGlyGlyAlaAlaThrGly 7222

RESULT 9
US-11-096-568A-20771
; Sequence 20771, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20771
;   LENGTH: 413
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(413)
;   OTHER INFORMATION: Ceres Seq. ID no. 12387083
US-11-096-568A-20771

Alignment Scores:
Pred. No.:              1.75e-06        Length:         413
Score:                  228.50          Matches:        134
Percent Similarity:     35.4%           Conservative:   41
Best Local Similarity:  27.1%           Mismatches:     174
Query Match:            7.4%            Indels:         145
DB:                     7               Gaps:           24

US-10-677-865A-2 (1-1590) x US-11-096-568A-20771 (1-413)

Qy         62 TGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACC 121

Db          5 CysPheLeu---ProSerCysSerSer------------ThrLeuLeuProSerProPro 19

Qy        122 CCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCCTGCCTCACTGGGAAGGGTG 181

Db         20 ProArgSerAsnGlnSerGlnSerArgProSerGlnArgArgAla---------ArgGlu 36

Qy        182 ACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTG 241

Db         37 AsnProHisSerProIleProThrProGlnPro---------------ProPhePro--- 50

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301

Db         51 ------------------------------------------------ProArg--- 52

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361

Db         53 ---------------LeuLeuGlnIleGlnSerSerPro---SerAlaAsnProProArg 66

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421

Db         67 Leu--Gln---IleProGlyArgGlnAlaAlaAsnGlnProTrpSerAlaSerArgArgS 85

Qy        422 CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCT 481

Db         85 erCysSerSerTrpAlaSerSerAlaSerThrAlaSerProArgArgSerSerGluAlaA 105

Qy        482 TTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACA 541

Db        105 rgThrAlaArgSerSer---------------ArgAlaSerArgProProSerArgCysP 120

Qy        542 ACTCTTAGCGC------GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTT 595

Db        120 roSerProArgPheSerSerCysThrSerProSerArgThrArgSerSerProThrSerT 140

Qy        596 CCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTA 655

Db        140 hrProThrThrArgProTrpThrHisProProArgAlaProThrArgSerAlaAlaCys- 159

Qy        656 GTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCT 715

Db        160 --------CysProAlaSerProProThrGlySerProCysSerSerSerSerProPro- 176

Qy        716 ACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACC 775

Db        177 -----ThrSerSerArgSerSerCysSerArgSerSerPro------------------- 188
```

```
Qy        776 AGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTC 835
                                                 ||||||||||
Db        189 ----------------------------SerProProProCysSer------------- 194

Qy        836 CAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGG 895
                  |||   ||||||   |||
Db        195 --------ProSerProProSerThrPro----------------------------- 201

Qy        896 TGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCT 955
                                           |||      |||   |||      :::
Db        202 ----------------------------ProSerAlaThrArgSerProSerArgAlaC 212

Qy        956 ATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCT 1015
                                                 ||||||     |||:::
Db        212 ys-------------------------------------SerProSerCysProAlaS 219

Qy       1016 TTGCAGCTGGG---CCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCC 1072
                  |||   ||||||   |||                ||||||:::|||      |
Db        219 erGlyGlyGlySerProProArgSerSerProProSerArgSerSerSerProThrThrP 239

Qy       1073 CCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCC 1132
                |||||      ::::::|||||||||
Db        239 roSerProCysSerSerSerSerAla-------------------------------- 247

Qy       1133 AGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCC------------TGCTCTCCCT 1180
                          |||      :::      |||            ||||||||||
Db        248 ----------SerSerSerProArgThrThrAlaProAlaSerArgAlaCysSerProS 264

Qy       1181 CCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATC 1240
                |||||||      |||:::   |||:::   |||:::||||||   |||        |
Db        264 erSerSerProSerProThrCysSerAlaSerSerThrSerAlaSerSerGlyThrSerP 284

Qy       1241 CCTGCGGCAGTGCTTCCCAGAGCCCCTGC------------------TCCCCACCAGGCA 1282
                ||            :::   |||||||||                |||||||||
Db        284 ro------------AlaSerSerProCysSerArgThrThrArgAlaSerProProCysA 300

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
                   :::   :::||||||      |||   ||||||   :::::::::
Db        300 laArgAlaArgThrSerSerGlyGlySer---SerSerProArgValLeu---------S 316

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
                   |||:::|||:::||||||      |||      |||   |||:::        :
Db        316 erSerSerArgSerThrSerSerSerProProSerSerSerAlaSerAlaProGlySerS 336

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
                ::|||   |||      :::|||               |||      |||   |||
Db        336 erGlyAspPro------AlaProArgArgGlnGlySerCysSerAlaSerSerCysSerP 354

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
                |||||||                |||||   |||
Db        354 roArgSerAlaAlaSerSerCysSerArgSerSerArg 366

RESULT 10
US-11-052-554A-157
; Sequence 157, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 157
;   LENGTH: 783
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-157

Alignment Scores:
Pred. No.:              3.2e-06     Length:          783
Score:                  223.50      Matches:         141
Percent Similarity:     30.4%       Conservative:    21
Best Local Similarity:  26.5%       Mismatches:      195
Query Match:            7.2%        Indels:          176
DB:                     7           Gaps:            21

US-10-677-865A-2 (1-1590) x US-11-052-554A-157 (1-783)

Qy       1564 GGGGTAGGAAAACTTCAGGGTCAGCTAGCTGGGGCCCCAGAGGCTTCA------------ 1517
                |||||||||   :::   |||      |||||||||      |||
Db        207 GlyValGlyGlyValGlyGlyLeuGlyAlaGlyAlaGlyGlyAlaGlyGlyValGlyGly 226

Qy       1516 ------------CTTGGGCTAGGATATCCCGGATGGAGCGGCAGGGGATCTTTCCAGCAC 1469
                         ||||||   |||      ||| ||||||   |||      :::   ||
Db        227 AlaGlyGlyLeuLeuGlyAlaGlyGlyHisGly-GlyAlaGlyGlyLeuGlyAlaValTh 246

Qy       1468 TGCTGGAGCCACAGGGCTTGGCACCAGCGGAGGGATCAGGATGGGGAGAGCCATCGGGGC 1409
                |   |||      |||   |||   :::||||||:::      |||          |||
Db        246 rGlyGlyValGlyGlyThrGlyGlyAlaGlyGlyLeuLeuAlaGlyLeuLeuAlaGly-- 265

Qy       1408 CCCCAGTCAGTGTCAAGGAGGAGACAGACATGCAAGGGTGACCAGAAGAGCTGGACTTGC 1349
                         ||||||                                  ||||||
Db        266 ------------ProGlyGly--------------------------AlaGlyGlyTh 272

Qy       1348 TGCCACAAGGCTGAAGGATGATTTTGCCACTCGATTGGGAACTGGAGCTGCTGCTGAAGG 1289
                   |||          |||                |||   |||   |||
Db        272 rGlyGlyArg---------GlyPheLeuAsnAsnGlyGlyValGlyGlyAlaGlyGlyAs 289

Qy       1288 AGCCGGTGCCTGGTGGGGAGCAGGGGCTCTGGGAAGCACTGCCGCAGGGATGGTAGGGTG 1229
                |||||      :::|||||||||      |||   :::              :::|||
Db        289 nAlaGlyLeuLeuPheGlyAlaGlyGlyThrGlyGlySerGlyGlyAlaGlyLeuGlyGl 309

Qy       1228 AACCGGCGCTGCTGGAAATGCT----------------------AGAACTGCTGGGGA 1193
                   |||   ||||||   |||                              ||||||
Db        309 yAspGlyGlyAlaGlyGlyAlaGlyGlyAsnThrGlyValLeuPheGlyAsnAlaGlySe 329

Qy       1192 CTCGAGAACTGGAGGGAGAGCAGGGTCCCTTGGAGCCCGTGGAGCCGCCTCCACAGAGCT 1133
                   |||||||||        :::   |||      |||   |||      :::
Db        329 rGlyGlyThrGlyGlyPheGlyAspThrAspGlyGlyAlaGlyGlyAlaGlyGlyAspAl 349

Qy       1132 GGACCCCACCAGTCCCCACTGGCTGGAACGCAATGGCCGAGGAAGCTGCCGACTGGCTGG 1073
                         |||||      |||   |||              |||||
Db        349 aGly--------------TrpLeuGlySerGlyGlyValGlyGly---------AlaGl 361

Qy       1072 GGATGATGGGGTTGCTGGAGAAGTATTTGCCCTCAGAGATGGGGGGCCCAGCTGCAAAGG 1013
                |      |||      |||                |||||||||
Db        361 yGlyPheGlyGluThrGly-----------------AspGlyGly-------------- 370

Qy       1012 AAGGGACCCCTGGAGAGCCTTTCACAGGGTTCTCTTTGGTGAAGTAGCCCACAGGATAGA 953

Db        370 ------------------------------------------------------------ 370

Qy        952 TTTTACCCTTACTGTAGGTCATGCCTGGAACCAGATAACTGTCAGAGGAGCCACCCACCA 893
                         ||||||   |||                      |||
Db        371 ------------ValGlyGlyAla--------------------GlyGlyLysAlaGl 379

Qy        892 CCTCGTAGCCACCATAGGATTTGTCTACAGAGGTGATTGGGGGACAGGGCTTGCCTGGAA 833
                  |||:::                        |||:::   |||   |||   |||
Db        379 yLeuLeuIle--------------------GlyAsnGlyGlyAlaGlyGlyAlaGlyGl 392

Qy        832 GGCCACCATTGCTACAGGGGGGACCTTGAACCACTCCAGGGGCACCAGAACCGTGCTGGT 773
                         ||||||||||||
Db        392 yGlnGlyAlaValThrGlyGlyThr---------------------------------- 400

Qy        772 CCACCACCACCACCACAGGCCTCTGACCCCCTGACACAGAGTGGGAGCTGGGGATGTAGG 713
                                                          |||||||||      ||
Db        401 ---------------------------------------GlyGlyAlaGlyGlyAspGl 407
```

```
Qy        712 GGCCAGAGTGCGAGACGATGGGCCCTCC--------------------ACTGCAGGGAG 674
Db        407 yValLeuIleGlyAsnGlyGlyAsnAlaGlyIleGlyGlyThrGlyProThrAlaGlyAs 427

Qy        673 AGTCGGGGATGTCCGAACTACAGGGACGCTGGTTGGAGCTGAC---------------- 631
Db        427 pThrGlyAlaGlyGlyIleSerGlyLeuLeuLeuGlyAlaAspGlyPheAsnThrProAl 447

Qy        630 ---------------------------------GCTTTGGCCACTGCTGGATACCCCAG 605
Db        447 aSerAlaSerProLeuHisThrLeuLysGlnGlnAlaLeuAlaAlaIleAsnAlaProTh 467

Qy        604 AGGTTTGGGAAGAGGAAGAGCTTTG---------TCCAGGCTGGGAAGGGTTTAGTATTC 554
Db        467 rGlnThrLeuThrGlyArgProLeuIleGlyAsnGlyThrProGlyAlaValGlySerGl 487

Qy        553 CGCGCTAAGAGTTGTCATTGGTTGGCAGAGCAGAGCCATTCCCTACTTGGAAGCTGCTGC 494
Db        487 yAlaThrGlyAlaProGlyGlyTrpLeu------------------------------- 496

Qy        493 TGCTGAACTGAAAGCTGCTGCTGCTGCTCGAATGAGAGCTGCTGCTTCCCGAGTGAGAGC 434
Db        497 -------LeuGlyAspGlyGlyAlaGlyGlySerGlyAlaAlaGlySerGlyAlaProGl 514

Qy        433 CGCTGCTTCCCGAGTGAGAGCTGCTGCTCCCCAGCTGGGAGGAACCGGATGCACCTTGTA 374
Db        514 yGlyAlaGlyGlyAla---AlaGlyLeuTrpGlyThrGlyGlyAlaGlyGlyAlaGlyGl 533

Qy        373 GAGTAGAGCCAGATCCGGAGGAGTAGCTGACCTGGGAATACCCCGTTCCTGGCTTAAAAG 314
Db        533 ySerSerAlaGlyGlyGlyGlyAlaGlyGlyAlaGlyGlyAlaGlyGlyTrpLeu----- 551

Qy        313 ATCCTGCAGAACCACCCTGGGCAATGCTGGATCCGCTGGAGCTACCACTGGAGCCACCAC 254
Db        552 ---------------LeuGlyAsp---GlyGlyAlaGlyGlyIleGlyGlyAlaSerTh 565

Qy        253 CAGAGCTTCTGGCACTGGAAATGGA--------------------GCTGCCAGAACTGC 215
Db        565 rValLeuGlyGlyThrGlyGlyGlyGlyValGlyGlyLeuTrpGlyAlaGlyGlyAl 585

Qy        214 TGGAGCCACTGTAGCTACTGAAGCCGCTGGAGTCACCCTTCCCAGTGAGGCAGGGGTCGT 155
Db        585 aGlyGlyAlaGlyGlyThrGlyLeuValGly-----------GlyAspGlyGlyAlaGl 601

Qy        154 TAGGGGAGGTGATACGCGTGGGGTCCTTACAGGGGTC 118
Db        601 yGlyAlaGlyGlyThrGlyGlyLeuLeuAlaGlyLeu 613

RESULT 11
US-11-169-232-14
; Sequence 14, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/169,232
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 14:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 941 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-11-169-232-14

Alignment Scores:
Pred. No.:              4.39e-06        Length:         941
Score:                  221.00          Matches:        135
Percent Similarity:     33.0%           Conservative:   40
Best Local Similarity:  25.5%           Mismatches:     252
Query Match:            7.1%            Indels:         103
DB:                     7               Gaps:           26

US-10-677-865A-2 (1-1590) x US-11-169-232-14 (1-941)

Qy         89 CCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCT 148
Db        124 ProThrSerLysValLeuAlaLysProThrProLysAlaGluThrThrThrLysGlyPro 143

Qy        149 CCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA--- 205
Db        144 AlaLeuThrThrProLysGluProThrProThrThrProLysGluProAlaSerThrThr 163

Qy        206 ------------GTGGCTCCAGCAGTTCTGGCAGCTCCATTT------------CCAGTG 241
Db        164 ProLysGluProThrProThrThrIleLysSerAlaProThrThrProLysGluProAla 183

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301
Db        184 ProThrThrThrLysSerAlaPro---ThrThrProLysGluProAla---ProThrThr 201

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361
Db        202 ThrLysGlu--------------------------ProAlaProThrThrProLysGlu 212

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421
Db        213 ProAlaProThrThrThrLysGluProAlaProThrThrThrLysSerAlaProThrThr 232

Qy        422 ---CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCA 478
Db        233 ProLysGluProAlaProThrThrProLysLysProAlaProThrThrProLysGluPro 252

Qy        479 GCTTTCAGTTCA------------------------------GCAGCAGCAGCTTCC 505
Db        253 AlaProThrThrProLysGluProThrProThrThrProLysGluProAlaProThrThr 272

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
```

```
Db        273 Lys--------------------GluProAlaProThrThrProLysGlu------Pro 283

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
Db        284 AlaProThrAlaProLysLys---ProAlaProThrThrProLysGluProAlaProThr 302

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
Db        303 ThrProLysGluProAlaProThrThrThrLysGluProSerProThrThrProLysGlu 322

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCC---------CCAGCTCCCACTCTG 736
Db        323 ProAlaProThrThrThrLysSerAlaProThrThrThrLysGluProAlaProThrThr 342

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
Db        343 ThrLys---------------------------------SerAlaProThrThrPro--- 350

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAA 856
Db        351 ---------LysGluProSerProThrThrThrLysGluProAlaProThrThrProLys 367

Qy        857 TCACCTCTGTAGACA---------AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCT 907
Db        368 GluProAlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaPro 387

Qy        908 CTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCT 967
Db        388 ThrThrPro------LysGluProAlaProThrThrThrLysLysProAlaProThrAla 405

Qy        968 ------------------ACTTCACCAAAGAGA---------ACCCTGTGAAAGGCTCTC 1000
Db        406 ProLysGluProAlaProThrThrProLysGluThrAlaProThrThrProLysLysLeu 425

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCC--------------------CCATCTCTG 1039
Db        426 ThrProThrThrProGluLysLeuAlaProThrThrProGluLysProAlaProThrThr 445

Qy       1040 AGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCA 1099
Db        446 ProGluGluLeuAlaProThrThrProGluGluProThrProThrThrProGluGluPro 465

Qy       1100 TTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCT 1159
Db        466 AlaProThrThrProLysAlaAlaAlaProAsnThrProLysGluProAlaProThrThr 485

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCA 1219
Db        486 ProLysGluProAla---ProThrThrProLysGluProAla------------ProThr 500

Qy       1220 GCGCCG------GTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA------------ 1261
Db        501 ThrProLysGluThrAlaProThrThrProLysGlyThrAlaProThrThrLeuLysGlu 520

Qy       1262 ---GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGA 1318
Db        521 ProAlaProThrThrProLysLysProAlaProLysGluLeuAlaProThrThrThrLys 540

Qy       1319 GTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCA 1378
Db        541 GluProThrSerThrThrSerAspLysProAla---ProThrThrProLysGlyThrAla 559

Qy       1379 TGTCTGTCTCCT------CCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTG 1432
Db        560 ProThrThrProLysGluPro--------------AlaPro---ThrThrProLysGlu 573

Qy       1433 ATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCT 1492
Db        574 ProAlaProThrThrProLysGlyThrAlaProThrThrLeu-LysGluProAlaProTh 593

Qy       1493 CCATCCGGGATATCCTAGCCCAAGTGAA 1520
```

```
Db        593 rThrProLysLysProAlaProLysGlu 602

RESULT 12
US-11-169-232-84
; Sequence 84, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/169,232
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 84:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1022 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-11-169-232-84

Alignment Scores:
Pred. No.:              4.36e-06      Length:          1022
Score:                  221.00        Matches:         135
Percent Similarity:     33.0%         Conservative:    40
Best Local Similarity:  25.5%         Mismatches:      252
Query Match:            7.1%          Indels:          103
DB:                     7             Gaps:            26

US-10-677-865A-2 (1-1590) x US-11-169-232-84 (1-1022)

Qy         89 CCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCT 148
Db        205 ProThrSerLysValLeuAlaLysProThrProLysAlaGluThrThrThrLysGlyPro 224

Qy        149 CCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA--- 205
```

```
Db        225 AlaLeuThrThrProLysGluProThrProThrThrProLysGluProAlaSerThrThr 244

Qy        206 ------------GTGGCTCCAGCAGTTCTGGCAGCTCCATTT------------CCAGTG 241
                          :::   :::||||||               |||
Db        245 ProLysGluProThrProThrThrIleLysSerAlaProThrThrProLysGluProAla 264

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301
              |||               ||||||        |||   :::||||||   |||
Db        265 ProThrThrThrLysSerAlaPro---ThrThrProLysGluProAla---ProThrThr 282

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361
                  :::                        |||         ||||||   :::
Db        283 ThrLysGlu-------------------------ProAlaProThrThrProLysGlu 293

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421
               |||                   |||   |||   :::      :::|||      |||
Db        294 ProAlaProThrThrThrLysGluProAlaProThrThrThrLysSerAlaProThrThr 313

Qy        422 ---CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCA 478
                  :::|||   |||       |||   :::       |||
Db        314 ProLysGluProAlaProThrThrProLysLysProAlaProThrThrProLysGluPro 333

Qy        479 GCTTTCAGTTCA--------------------------------GCAGCAGCAGCTTCC 505
              |||   ::::::                              |||      :::
Db        334 AlaProThrThrProLysGluProThrProThrThrProLysGluProAlaProThrThr 353

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
              |||                   :::|||      |||         |||
Db        354 Lys--------------------GluProAlaProThrThrProLysGlu------Pro 364

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
               |||:::     |||     |||   |||                |||||
Db        365 AlaProThrAlaProLysLys---ProAlaProThrThrProLysGluProAlaProThr 383

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
                     |||                              |||||||||   |||
Db        384 ThrProLysGluProAlaProThrThrThrLysGluProSerProThrThrProLysGlu 403

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCC---------CCAGCTCCCACTCTG 736
                 |||::::::         |||||||||:::         ||||||||||||
Db        404 ProAlaProThrThrThrLysSerAlaProThrThrThrLysGluProAlaProThrThr 423

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
                 :::                              |||                |||
Db        424 ThrLys---------------------------------SerAlaProThrThrPro--- 431

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAA 856
                    |||   |||                                 |||         |||:::
Db        432 ---------LysGluProSerProThrThrThrLysGluProAlaProThrThrProLys 448

Qy        857 TCACCTCTGTAGACA---------AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCT 907
               |||      |||            |||         |||             ||||||
Db        449 GluProAlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaPro 468

Qy        908 CTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCT 967
               |||          :::     ||||||        |||             |||
Db        469 ThrThrPro------LysGluProAlaProThrThrThrLysLysProAlaProThrAla 486

Qy        968 -----------------ACTTCACCAAAGAGA---------ACCCTGTGAAAGGCTCTC 1000
                              |||:::||||||            |||    |||   |||
Db        487 ProLysGluProAlaProThrThrProLysGluThrAlaProThrThrProLysLysLeu 506

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCC--------------------CCATCTCTG 1039
                  :::   |||   :::|||   |||                        |||:::
Db        507 ThrProThrThrProGluLysLeuAlaProThrThrProGluLysProAlaProThrThr 526

Qy       1040 AGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCA 1099
                        :::|||   ||||||      |||                   |||
Db        527 ProGluGluLeuAlaProThrThrProGluGluProThrProThrThrProGluGluPro 546

Qy       1100 TTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCT 1159
                  ::::::                  ::::::      |||   ||||||
Db        547 AlaProThrThrProLysAlaAlaAlaProAsnThrProLysGluProAlaProThrThr 566

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCA 1219
              |||::::::||||||   |||         :::   ||||||            |||
Db        567 ProLysGluProAla---ProThrThrProLysGluProAla------------ProThr 581

Qy       1220 GCGCCG------GTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA------------ 1261
                 |||            ||||||   |||            |||
Db        582 ThrProLysGluThrAlaProThrThrProLysGlyThrAlaProThrThrLeuLysGlu 601

Qy       1262 ---GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGA 1318
                 ||||||         :::   |||||||||         ||||||         :::
Db        602 ProAlaProThrThrProLysLysProAlaProProLysGluLeuAlaProThrThrThrLys 621

Qy       1319 GTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCA 1378
                    |||:::   |||         |||   |||                  |||
Db        622 GluProThrSerThrThrSerAspLysProAla---ProThrThrProLysGlyThrAla 640

Qy       1379 TGTCTGTCTCCT------CCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTG 1432
                    :::|||      |||                 ||||||         |||
Db        641 ProThrThrProLysGluPro--------------AlaPro---ThrThrProLysGlu 654

Qy       1433 ATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCT 1492
                    |||       |||         ||||||      ||| |||:::|||   |||
Db        655 ProAlaProThrThrProLysGlyThrAlaProThrThrLeu-LysGluProAlaProTh 674

Qy       1493 CCATCCGGGATATCCTAGCCCAAGTGAA 1520
                    |||      |||:::|||   |||
Db        674 rThrProLysLysProAlaProLysGlu 683
```

```
RESULT 13
US-11-169-232-74
; Sequence 74, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/169,232
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
```

```
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 74:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 1038 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-11-169-232-74

Alignment Scores:
Pred. No.:              4.35e-06      Length:         1038
Score:                  221.00        Matches:        135
Percent Similarity:     33.0%         Conservative:   40
Best Local Similarity:  25.5%         Mismatches:     252
Query Match:            7.1%          Indels:         103
DB:                     7             Gaps:           26

US-10-677-865A-2 (1-1590) x US-11-169-232-74 (1-1038)

Qy         89 CCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCT 148
Db        189 ProThrSerLysValLeuAlaLysProThrProLysAlaGluThrThrThrLysGlyPro 208

Qy        149 CCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA--- 205
Db        209 AlaLeuThrThrProLysGluProThrProThrThrProLysGluProAlaSerThrThr 228

Qy        206 -----------GTGGCTCCAGCAGTTCTGGCAGCTCCATTT-----------CCAGTG 241
Db        229 ProLysGluProThrProThrThrIleLysSerAlaProThrThrProLysGluProAla 248

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301
Db        249 ProThrThrThrLysSerAlaPro---ThrThrProLysGluProAla---ProThrThr 266

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361
Db        267 ThrLysGlu--------------------------ProAlaProThrThrProLysGlu 277

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421
Db        278 ProAlaProThrThrThrLysGluProAlaProThrThrThrLysSerAlaProThrThr 297

Qy        422 ---CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCA 478
Db        298 ProLysGluProAlaProThrThrProLysLysProAlaProThrThrProLysGluPro 317

Qy        479 GCTTTCAGTTCA--------------------------------GCAGCAGCAGCTTCC 505
Db        318 AlaProThrThrProLysGluProThrProThrThrProLysGluProAlaProThrThr 337

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
Db        338 Lys--------------------GluProAlaProThrThrProLysGlu------Pro 348

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
Db        349 AlaProThrAlaProLysLys---ProAlaProThrThrProLysGluProAlaProThr 367

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
Db        368 ThrProLysGluProAlaProThrThrThrLysGluProSerProThrThrProLysGlu 387

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCC---------CCAGCTCCCACTCTG 736
Db        388 ProAlaProThrThrThrLysSerAlaProThrThrThrLysGluProAlaProThrThr 407

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
Db        408 ThrLys---------------------------------SerAlaProThrThrPro--- 415

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAA 856
Db        416 ---------LysGluProSerProThrThrThrLysGluProAlaProThrThrProLys 432

Qy        857 TCACCTCTGTAGACA---------AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCT 907
Db        433 GluProAlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaPro 452

Qy        908 CTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCT 967
Db        453 ThrThrPro------LysGluProAlaProThrThrThrLysLysProAlaProThrAla 470

Qy        968 -----------------ACTTCACCAAAGAGA---------ACCCTGTGAAAGGCTCTC 1000
Db        471 ProLysGluProAlaProThrThrProLysGluThrAlaProThrThrProLysLysLeu 490

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCC---------------------CCATCTCTG 1039
Db        491 ThrProThrThrProGluLysLeuAlaProThrThrProGluLysProAlaProThrThr 510

Qy       1040 AGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCA 1099
Db        511 ProGluGluLeuAlaProThrThrProGluGluProThrProThrThrProGluGluPro 530

Qy       1100 TTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCT 1159
Db        531 AlaProThrThrProLysAlaAlaAlaProAsnThrProLysGluProAlaProThrThr 550

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCA 1219
Db        551 ProLysGluProAla---ProThrThrProLysGluProAla------------ProThr 565

Qy       1220 GCGCCG------GTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA------------ 1261
Db        566 ThrProLysGluThrAlaProThrThrProLysGlyThrAlaProThrThrLeuLysGlu 585

Qy       1262 ---GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGA 1318
Db        586 ProAlaProThrThrProLysLysProAlaProLysGluLeuAlaProThrThrThrLys 605

Qy       1319 GTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCA 1378
Db        606 GluProThrSerThrThrSerAspLysProAla---ProThrThrProLysGlyThrAla 624

Qy       1379 TGTCTGTCTCCT------CCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTG 1432
Db        625 ProThrThrProLysGluPro--------------AlaPro---ThrThrProLysGlu 638

Qy       1433 ATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCT 1492
Db        639 ProAlaProThrThrProLysGlyThrAlaProThrThrLeu-LysGluProAlaProTh 658

Qy       1493 CCATCCGGGATATCCTAGCCCAAGTGAA 1520
Db        658 rThrProLysLysProAlaProLysGlu 667

RESULT 14
US-11-169-232-58
; Sequence 58, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
```

```
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/11/169,232
;             FILING DATE: 28-Jun-2005
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;    INFORMATION FOR SEQ ID NO: 58:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1049 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-11-169-232-58

Alignment Scores:
Pred. No.:                4.35e-06      Length:          1049
Score:                    221.00        Matches:         135
Percent Similarity:       33.0%         Conservative:    40
Best Local Similarity:    25.5%         Mismatches:      252
Query Match:              7.1%          Indels:          103
DB:                       7             Gaps:            26

US-10-677-865A-2 (1-1590) x US-11-169-232-58 (1-1049)

Qy         89 CCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCT 148
              |||      :::   ||||||                 |||         |||
Db        232 ProThrSerLysValLeuAlaLysProThrProLysAlaGluThrThrThrLysGlyPro 251

Qy        149 CCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA--- 205
              ||||||||||||                ||||||           :::|||
Db        252 AlaLeuThrThrProLysGluProThrProThrThrProLysGluProAlaSerThrThr 271

Qy        206 ------------GTGGCTCCAGCAGTTCTGGCAGCTCCATTT------------CCAGTG 241
                          :::   :::||||||                |||
Db        272 ProLysGluProThrProThrThrIleLysSerAlaProThrThrProLysGluProAla 291

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301
              |||           ||||||        |||   :::||||||   |||
Db        292 ProThrThrThrLysSerAlaPro---ThrThrProLysGluProAla---ProThrThr 309

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361
                  :::                          |||         ||||||   :::
Db        310 ThrLysGlu--------------------------ProAlaProThrThrProLysGlu 320

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421
              |||                  |||   |||   :::      :::|||      |||
Db        321 ProAlaProThrThrThrLysGluProAlaProThrThrThrLysSerAlaProThrThr 340

Qy        422 ---CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCA 478
                    :::|||   |||      |||   :::      |||
Db        341 ProLysGluProAlaProThrThrProLysLysProAlaProThrThrProLysGluPro 360

Qy        479 GCTTTCAGTTCA--------------------------------GCAGCAGCAGCTTCC 505
              |||   ::::::                                |||      :::
Db        361 AlaProThrThrProLysGluProThrProThrThrProLysGluProAlaProThrThr 380

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
              |||                    :::|||      |||      |||
Db        381 Lys--------------------GluProAlaProThrThrProLysGlu------Pro 391

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
                 |||:::      |||      |||   |||                |||||
Db        392 AlaProThrAlaProLysLys---ProAlaProThrThrProLysGluProAlaProThr 410

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
                       |||                          |||||||||   |||
Db        411 ThrProLysGluProAlaProThrThrThrLysGluProSerProThrThrProLysGlu 430

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCC---------CCAGCTCCCACTCTG 736
                    |||::::::          ||||||||:::         ||||||||||
Db        431 ProAlaProThrThrThrLysSerAlaProThrThrThrLysGluProAlaProThrThr 450

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
                 :::                                  |||           |||
Db        451 ThrLys---------------------------------SerAlaProThrThrPro--- 458

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAA 856
                       |||  |||                        |||        |||:::
Db        459 ---------LysGluProSerProThrThrThrLysGluProAlaProThrProThrPro 475

Qy        857 TCACCTCTGTAGACA---------AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCT 907
                 |||      |||         |||      |||      |||          ||||||
Db        476 GluProAlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaPro 495

Qy        908 CTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCT 967
                 |||           :::      ||||||         |||           |||
Db        496 ThrThrPro------LysGluProAlaProThrThrThrLysLysProAlaProThrAla 513

Qy        968 ------------------ACTTCACCAAAGAGA---------ACCCTGTGAAAGGCTCTC 1000
                               |||:::||||||           |||      |||   |||
Db        514 ProLysGluProAlaProThrThrProLysGluThrAlaProThrThrProLysLysLeu 533

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCC--------------------CCATCTCTG 1039
                  :::   |||   :::|||   |||                     |||:::
Db        534 ThrProThrThrProGluLysLeuAlaProThrThrProGluLysProAlaProThrThr 553

Qy       1040 AGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCA 1099
                    :::|||   ||||||      |||                        |||
Db        554 ProGluGluLeuAlaProThrThrProGluGluProThrProThrThrProGluGluPro 573

Qy       1100 TTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCT 1159
                    ::::::                ::::::      |||   ||||||
Db        574 AlaProThrThrProLysAlaAlaAlaProAsnThrProLysGluProAlaProThrThr 593

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCA 1219
              |||::::::||||||   |||          :::   |||||             |||
Db        594 ProLysGluProAla---ProThrThrProLysGluProAla------------ProThr 608

Qy       1220 GCGCCG------GTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA------------ 1261
                 |||           ||||||   |||          |||
Db        609 ThrProLysGluThrAlaProThrThrProLysGlyThrAlaProThrThrLeuLysGlu 628

Qy       1262 ---GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGA 1318
                 ||||||      :::   |||||||||          ||||||          :::
Db        629 ProAlaProThrThrProLysLysProAlaProLysGluLeuAlaProThrThrThrLys 648

Qy       1319 GTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCA 1378
                    |||:::   |||          |||   |||                |||
Db        649 GluProThrSerThrThrSerAspLysProAla---ProThrThrProLysGlyThrAla 667

Qy       1379 TGTCTGTCTCCT------CCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTG 1432
                 :::|||      |||           ||||||         |||
```

```
Db        668 ProThrThrProLysGluPro--------------AlaPro---ThrThrProLysGlu 681

Qy       1433 ATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCT 1492
                 |||      |||       ||||||     ||| |||:::|||   |||
Db        682 ProAlaProThrThrProLysGlyThrAlaProThrThrLeu-LysGluProAlaProTh 701

Qy       1493 CCATCCGGGATATCCTAGCCCAAGTGAA 1520
                 |||      |||:::|||   |||
Db        701 rThrProLysLysProAlaProLysGlu 710

RESULT 15
US-11-169-232-104
; Sequence 104, Application US/11169232
; Publication No. US20060025570A1
;   GENERAL INFORMATION:
;       APPLICANT: Turner, Katherine
;                  Clark, Stephen C.
;                  Jacobs, Kenneth
;                  Hewick, Rodney M.
;                  Gesner, Thomas G.
;       TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;       NUMBER OF SEQUENCES: 143
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/11/169,232
;            FILING DATE: 28-Jun-2005
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/10/124,557
;            FILING DATE: 16-Apr-2002
;            APPLICATION NUMBER: US 07/643,502
;            FILING DATE: 18-JAN-1991
;            APPLICATION NUMBER: US 07/546,114
;            FILING DATE: 29-JUN-1990
;            APPLICATION NUMBER: US 07/457,196
;            FILING DATE: 29-DEC-1989
;            APPLICATION NUMBER: US 07/390,901
;            FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 104:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1140 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-11-169-232-104

Alignment Scores:
Pred. No.:              4.31e-06      Length:          1140
Score:                  221.00        Matches:          135
Percent Similarity:     33.0%         Conservative:      40
Best Local Similarity:  25.5%         Mismatches:       252
Query Match:            7.1%          Indels:           103
DB:                     7             Gaps:              26
```

```
US-10-677-865A-2 (1-1590) x US-11-169-232-104 (1-1140)

Qy         89 CCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCT 148
              |||      :::   ||||||                 |||                 |||
Db        323 ProThrSerLysValLeuAlaLysProThrProLysAlaGluThrThrThrLysGlyPro 342

Qy        149 CCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA--- 205
              ||||||||||||                ||||||              :::|||
Db        343 AlaLeuThrThrProLysGluProThrProThrThrProLysGluProAlaSerThrThr 362

Qy        206 -----------GTGGCTCCAGCAGTTCTGGCAGCTCCATTT-----------CCAGTG 241
                          :::   :::||||||               |||
Db        363 ProLysGluProThrProThrThrIleLysSerAlaProThrThrProLysGluProAla 382

Qy        242 CCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTG 301
              |||                |||||       |||   :::||||||   |||
Db        383 ProThrThrThrLysSerAlaPro---ThrThrProLysGluProAla---ProThrThr 400

Qy        302 GTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGAT 361
                  :::                          |||         ||||||   :::
Db        401 ThrLysGlu--------------------------ProAlaProThrThrProLysGlu 411

Qy        362 CTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACT 421
              |||                   |||  |||   :::      :::|||      |||
Db        412 ProAlaProThrThrThrLysGluProAlaProThrThrThrLysSerAlaProThrThr 431

Qy        422 ---CGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCA 478
                 :::|||   |||      |||  :::     |||
Db        432 ProLysGluProAlaProThrThrProLysLysProAlaProThrThrProLysGluPro 451

Qy        479 GCTTTCAGTTCA--------------------------------GCAGCAGCAGCTTCC 505
              |||   ::::::                              |||      :::
Db        452 AlaProThrThrProLysGluProThrProThrThrProLysGluProAlaProThrThr 471

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
              |||                  :::|||      |||         |||
Db        472 Lys--------------------GluProAlaProThrThrProLysGlu------Pro 482

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
              |||:::      |||     |||  |||                 ||||||
Db        483 AlaProThrAlaProLysLys---ProAlaProThrThrProLysGluProAlaProThr 501

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
                    |||                        |||||||||   |||
Db        502 ThrProLysGluProAlaProThrThrThrLysGluProSerProThrThrProLysGlu 521

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCC---------CCAGCTCCCACTCTG 736
                  |||::::::          |||||||||:::         ||||||||||||
Db        522 ProAlaProThrThrThrLysSerAlaProThrThrThrLysGluProAlaProThrThr 541

Qy        737 TGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTG 796
                 :::                             |||                 |||
Db        542 ThrLys--------------------------------SerAlaProThrThrPro--- 549

Qy        797 GAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAA 856
                    |||   |||                          |||         |||:::
Db        550 ---------LysGluProSerProThrThrThrLysGluProAlaProThrThrProLys 566

Qy        857 TCACCTCTGTAGACA---------AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCT 907
              |||      |||           |||          |||             ||||||
Db        567 GluProAlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaPro 586

Qy        908 CTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCT 967
              |||           :::      ||||||       |||                 |||
Db        587 ThrThrPro------LysGluProAlaProThrThrThrLysLysProAlaProThrAla 604

Qy        968 -----------------ACTTCACCAAAGAGA---------ACCCTGTGAAAGGCTCTC 1000
                              |||:::||||||           |||     |||   |||
Db        605 ProLysGluProAlaProThrThrProLysGluThrAlaProThrThrProLysLysLeu 624

Qy       1001 CAGGGGTCCCTTCCTTTGCAGCTGGGCCCC--------------------CCATCTCTG 1039
                  :::   |||   :::|||   |||                        |||:::
Db        625 ThrProThrThrProGluLysLeuAlaProThrThrProGluLysProAlaProThrThr 644
```

```
Qy       1040 AGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCA 1099
                     :::|||  ||||||     |||                          |||
Db        645 ProGluGluLeuAlaProThrThrProGluGluProThrProThrThrProGluGluPro 664

Qy       1100 TTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCT 1159
                :::::                    ::::::      |||   ||||||
Db        665 AlaProThrThrProLysAlaAlaAlaProAsnThrProLysGluProAlaProThrThr 684

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCA 1219
              |||::::::||||||   |||         :::   ||||||                |||
Db        685 ProLysGluProAla---ProThrThrProLysGluProAla----------ProThr 699

Qy       1220 GCGCCG------GTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGA------------ 1261
                |||           ||||||   |||           |||
Db        700 ThrProLysGluThrAlaProThrThrProLysGlyThrAlaProThrThrLeuLysGlu 719

Qy       1262 ---GCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGA 1318
                 ||||||         :::   |||||||||         ||||||         :::
Db        720 ProAlaProThrThrProLysLysProAlaProLysGluLeuAlaProThrThrThrLys 739

Qy       1319 GTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCA 1378
                  |||:::   |||         |||   |||                    |||
Db        740 GluProThrSerThrThrSerAspLysProAla---ProThrThrProLysGlyThrAla 758

Qy       1379 TGTCTGTCTCCT------CCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTG 1432
                :::|||     |||               ||||||        |||
Db        759 ProThrThrProLysGluPro--------------AlaPro---ThrThrProLysGlu 772

Qy       1433 ATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCT 1492
                  |||      |||         ||||||     ||| |||:::|||   |||
Db        773 ProAlaProThrThrProLysGlyThrAlaProThrThrLeu-LysGluProAlaProTh 792

Qy       1493 CCATCCGGGATATCCTAGCCCAAGTGAA 1520
                 |||      |||:::|||   |||
Db        792 rThrProLysLysProAlaProLysGlu 801
```

Search completed: March 21, 2006, 22:14:54
Job time : 57 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:06:39 ; Search time 6.5 Seconds
(without alignments)
4044.749 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 3102
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152023_26177/app_query.fasta_1
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss05h -USER=US10677865_@CGN_1_1_33_@runat_21032006_152023_26177
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2738	88.3	529	2	US-09-381-656-1	Sequence 1, Appli
2	2683	86.5	529	2	US-10-164-230-61	Sequence 61, Appl
3	282.5	9.1	78	2	US-09-641-931F-39	Sequence 39, Appl
4	263	8.5	258	2	US-10-104-047-3034	Sequence 3034, Ap
5	261.5	8.4	1367	2	US-09-487-558B-108	Sequence 108, App
6	256	8.3	386	2	US-09-461-774-10	Sequence 10, Appl
7	251.5	8.1	2972	2	US-09-579-181-2	Sequence 2, Appli
8	251.5	8.1	3118	2	US-09-579-181-1	Sequence 1, Appli
9	247.5	8.0	493	2	US-08-556-978B-59	Sequence 59, Appl
10	247.5	8.0	1566	2	US-09-581-472B-2	Sequence 2, Appli
11	247.5	8.0	1581	2	US-09-110-517-2	Sequence 2, Appli
12	245	7.9	595	1	US-08-425-069-4	Sequence 4, Appli
13	245	7.9	595	1	US-08-317-844B-4	Sequence 4, Appli
14	243.5	7.8	529	2	US-09-247-806-2	Sequence 2, Appli
15	243.5	7.8	529	2	US-09-861-597-2	Sequence 2, Appli
16	242	7.8	1259	2	US-09-949-016-10366	Sequence 10366, A
17	240	7.7	316	2	US-09-538-092-997	Sequence 997, App
18	234.5	7.6	343	2	US-10-148-306-21	Sequence 21, Appl
19	234.5	7.6	461	2	US-09-248-796A-23039	Sequence 23039, A
20	234.5	7.6	1970	2	US-09-538-092-1005	Sequence 1005, Ap
21	229	7.4	682	1	US-08-642-255-126	Sequence 126, App
22	229	7.4	682	1	US-08-397-633A-36	Sequence 36, Appl
c 23	225.5	7.3	908	2	US-08-714-741-44	Sequence 44, Appl
c 24	225	7.3	2088	2	US-09-548-372D-13	Sequence 13, Appl
c 25	225	7.3	2088	2	US-09-548-367D-13	Sequence 13, Appl
c 26	225	7.3	2088	2	US-09-551-853D-13	Sequence 13, Appl
c 27	225	7.3	2088	2	US-09-548-376D-13	Sequence 13, Appl
c 28	225	7.3	2088	2	US-09-548-373D-13	Sequence 13, Appl
c 29	225	7.3	2088	2	US-09-548-366F-13	Sequence 13, Appl
c 30	225	7.3	2088	2	US-09-548-368D-13	Sequence 13, Appl
31	224.5	7.2	1185	2	US-09-041-886-23	Sequence 23, Appl
32	224.5	7.2	1185	2	US-09-538-092-1209	Sequence 1209, Ap
33	222.5	7.2	100	2	US-09-411-067C-5	Sequence 5, Appli
34	222.5	7.2	1042	2	US-09-252-991A-30444	Sequence 30444, A
35	222.5	7.2	1184	2	US-09-266-225D-18	Sequence 18, Appl
36	221	7.1	941	2	US-07-757-022B-14	Sequence 14, Appl
37	221	7.1	1022	2	US-07-757-022B-84	Sequence 84, Appl
38	221	7.1	1038	2	US-07-757-022B-74	Sequence 74, Appl
39	221	7.1	1049	2	US-07-757-022B-58	Sequence 58, Appl
40	221	7.1	1140	2	US-07-757-022B-104	Sequence 104, App
41	221	7.1	1270	2	US-07-757-022B-44	Sequence 44, Appl
42	221	7.1	1311	2	US-07-757-022B-42	Sequence 42, Appl
43	221	7.1	1313	2	US-07-757-022B-142	Sequence 142, App
44	221	7.1	1314	2	US-07-757-022B-50	Sequence 50, Appl
45	221	7.1	1320	2	US-07-757-022B-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Alignment Scores:

Pred. No.:	1.03e-182	Length:	529
Score:	2738.00	Matches:	527
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	88.3%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-381-656-1 (1-529)

Qy 1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

```
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db        141 SerGlySerSerGlySerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db        201 ThrSerGlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db        281 LysProCysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db        401 SerSerSerIleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
Db        521 ProGlnGlyGluLeuLeuAspSerPro 529
```

RESULT 2
US-10-164-230-61
; Sequence 61, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 143
; OTHER INFORMATION: Xaa = Ser or Asn
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 151
; OTHER INFORMATION: Xaa = Ser or Cys
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 152
; OTHER INFORMATION: Xaa = Ser or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 202
; OTHER INFORMATION: Xaa = Ser or Phe
; FEATURE:
; NAME/KEY: VARIANT

; LOCATION: 282
; OTHER INFORMATION: Xaa = Pro or Thr
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 401
; OTHER INFORMATION: Xaa = Ser or Gly
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 408
; OTHER INFORMATION: Xaa = Ala or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 409
; OTHER INFORMATION: Xaa = Gly or Val
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 410
; OTHER INFORMATION: Xaa = Ser or Leu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 453
; OTHER INFORMATION: Xaa = Ser or Asn
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 527
; OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

Alignment Scores:

Pred. No.:	6.98e-179	Length:	529
Score:	2683.00	Matches:	516
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	97.5%	Mismatches:	12
Query Match:	86.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-164-230-61 (1-529)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SerGlySerSerLeuGlnGlyAlaSerGlySerSerGlnLeuGlySerSerSerSerHis 140

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||  ||||||||||||||||||||      ||||||||||||||||||||||||
Db        141 SerGly***SerGlySerHisSerGlySer******SerHisSerSerSerSerSerSer 160

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 PheGlnPheSerSerSerSerPheGlnValGlyAsnGlySerAlaLeuProThrAsnAsp 180

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AsnSerTyrArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGln 200

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 Thr***GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSer 220

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AspIleProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIle 240

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ProSerSerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHis 260

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlySerGlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGly 280

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 Lys***CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGly 300

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GlySerSerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrPro 320

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValGlyTyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAla 340

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGln 360

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SerAlaAlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCys 380

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyGlyGlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        401 ***SerSerIleSerSerSer*********ProTyrHisProCysGlySerAlaSerGln 420

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SerProCysSerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSer 440

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        441 GlyLysIleIleLeuGlnProCysGlySerLysSer***SerSerGlyHisProCysMet 460

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerValSerSerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSer 480

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaGlyAlaLysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArg 500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspIleLeuAlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeu 520

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||  ||||||
Db        521 ProGlnGlyGluLeuLeu***SerPro 529
```

```
RESULT 3
US-09-641-931F-39
; Sequence 39, Application US/09641931F
; Patent No. 6830885
; GENERAL INFORMATION:
;  APPLICANT: Lanctot, et al.
;  TITLE OF INVENTION: Nucleic Acid Molecule, Method and Kit for Selecting a
;  TITLE OF INVENTION:  Nucleic Acid Having A Desired Feature
;  FILE REFERENCE: 2003390-0001
;  CURRENT APPLICATION NUMBER: US/09/641,931F
;  CURRENT FILING DATE:  2003-04-08
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-641-931F-39

Alignment Scores:
Pred. No.:              6.34e-12        Length:         78
Score:                  282.50          Matches:        56
Percent Similarity:     81.8%           Conservative:   7
Best Local Similarity:  72.7%           Mismatches:     13
Query Match:            9.1%            Indels:         1
DB:                     2               Gaps:           1

US-10-677-865A-2 (1-1590) x US-09-641-931F-39 (1-78)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||:::||||||   ||||||   ||||||||||||||:::|||||||||
Db          1 MetGlySerSerGlnAlaProArgMetGlySerValGlyGlyHisGlyLeuMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              |||:::||||||   :::|||||||||   ||||||||||||||||||||   ||||||
Db         21 LeuMetAlaGly---IleLeuProGlyIleLeuAlaLysSerIleGlyThrLeuSerAsp 39

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         40 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuIleGlyLysThr 59

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231
                 |||:::      ||||||      ||||||||||||      ::::::
Db         60 GlySerAsnSerIleSerSerGlnGlyGlySerSerSerArgSerAlaAla 76

RESULT 4
US-10-104-047-3034
; Sequence 3034, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. 6943241el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3034
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-3034

Alignment Scores:
Pred. No.:              1.95e-10        Length:         258
Score:                  263.00          Matches:        94
Percent Similarity:     43.7%           Conservative:   20
Best Local Similarity:  36.0%           Mismatches:     113
Query Match:            8.5%            Indels:         34
DB:                     2               Gaps:           5

US-10-677-865A-2 (1-1590) x US-10-104-047-3034 (1-258)

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
              |||      |||:::      :::||||||      ||||||                |||
Db         15 SerThrLysCysArgCysGlySerArgIleAlaGlyProAsn---------AlaLeuGly 31

Qy        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCC---------- 213
                 |||    |||          ||||||   |||   |||
Db         32 SerGlyGlySerArgSerSerSerSerSerArgSerIleLeuSerSerSerIleLeu 51

Qy        214 ------------AGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGC 261
                         ||||||   ||||||   ||||||      ||||||
Db         52 SerSerSerIleProSerSerSerSerSerSerSerSerProSerSerSerHisSerSer 71

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAG 321
              ||||||   |||||||||   ||||||   :::      |||:::   :::
Db         72 SerSerProSerSerSerHisSerSerSerSerProSerSerSerSerSerThrSerSer 91

Qy        322 CCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGT 381
              |||   :::      |||      |||   ||||||   |||   |||
Db         92 ProSerSerSerSerSerSerSerSerSerSerProSerSerSerAsnSerSerSer 111

Qy        382 GCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCAC 441
              :::|||   ||||||          ||||||||||||   |||   ||||||   |||
Db        112 SerSerSerSerSerSerProSerSerSerSerSerSerSerSerSerSerProSerSer 131

Qy        442 TCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGC 501
              |||   ||||||   |||   ||||||||||||||||||          ||||||||||||
Db        132 SerSerSerSerProSerSerSerSerSerSerSerSerSerSerProSerSerSerSer 151

Qy        502 TTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTA 561
                          :::   |||:::      ::::::   :::|||
Db        152 SerSerProSerSerSerSerSerSerSerSerSerSerSerSerProSerSerSer 171

Qy        562 AACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGT 621
              :::||||||      |||   |||   |||||||||   :::|||      |||||||||
Db        172 SerProSerSerSerGlySerSerProSerSerSerAsnSerSerProSerSerSerSer 191

Qy        622 GGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGC 681
                        ||||||:::      |||                      ||||||
Db        192 SerSerProSerSerSerSerSerSerPro--------------------SerProArg 204

Qy        682 AGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCA 741
              |||      |||      |||   |||            ||||||:::   |||   |||
Db        205 SerSerSerProSerSerSerSerSerThrSerProSerThrSerSerProSer 224

Qy        742 GGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTG 801
                                                        |||   :::|||
Db        225 Ser--------------------------SerSerProSerSerSerSerPro------ 233

Qy        802 GTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACC 861
                          |||   :::      |||      :::|||      |||   :::
Db        234 ------SerSerSerCysProSerAlaAlaLeuGlyArgArgProGlnSerProGlnSer 251

Qy        862 TCT 864
              |||
Db        252 Ser 252

RESULT 5
US-09-487-558B-108
; Sequence 108, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. 6949356man, Thea
;  APPLICANT:  Royer, John
;  APPLICANT:  Salama, Sofie
```

```
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.130
;  CURRENT APPLICATION NUMBER: US/09/487,558B
;  CURRENT FILING DATE:  2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/487,558
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 446
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
;   LENGTH: 1367
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-108

Alignment Scores:
Pred. No.:                3.78e-10      Length:          1367
Score:                    261.50        Matches:         148
Percent Similarity:       39.1%         Conservative:    71
Best Local Similarity:    26.4%         Mismatches:      250
Query Match:              8.4%          Indels:          92
DB:                       2             Gaps:            19

US-10-677-865A-2 (1-1590) x US-09-487-558B-108 (1-1367)

Qy        119 ACCCCTGTAAGGACCCCA-----------------CGCGTATCACCTCCCCTAACGACC 160
Db        312 ThrProValProThrProSerSerSerThrThrGluSerSerSerAlaProValProThr 331

Qy        161 CCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTT 220
Db        332 ProSerSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSer 351

Qy        221 CTGGCAGCTCCATTTCCAGTGCCAGA-------------------------------- 246
Db        352 SerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaProVal 371

Qy        247 ---AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
Db        372 ThrSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSerSer 391

Qy        304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
Db        392 SerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaProValThr 411

Qy        364 GGCTCTACTCTACAAGGTGCATCCGGT---------TCCTCCCAGCTGGGGAGCAGCAGC 414
Db        412 SerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSerSerSer 431

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGC------AGCTCTCATTCGAGC 468
Db        432 AlaProValThrSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThr 451

Qy        469 AGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTG 528
Db        452 GluSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAla 471

Qy        529 CCA------------ACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCT 576
Db        472 ProValThrSerSerThrThrGluSerSerSerAlaProValProThrProSerSerSer 491

Qy        577 GGACAAAGCTCTTCCTCT------------TCCCAAACCTCTGGGGTATCCAGCAGTGGC 624
Db        492 ThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSerSerSerAlaPro 511

Qy        625 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC-------------- 669
Db        512 ValProThrProSerSerSerThrThrGluSerSerSerAlaProAlaProThrProSer 531

Qy        670 --------------------GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCT 708
Db        532 SerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSerSerSer 551

Qy        709 GGCCCCTAC---ATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTG 765
Db        552 AlaProValProThrProSerSerSerThrThrGluSerSerSerThrProValThrSer 571

Qy        766 GTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAAT 825
Db        572 SerThrThrGlu---SerSerSerAlaPro------ValProThrProSerSerSerThr 588

Qy        826 GGTGGCCTTCCAGGCAAGCCCTGTCCC---CCAATCACCTCTGTAGACAAATCCTATGGT 882
Db        589 ThrGluSerSerSerAlaProValProThrProSerSerSerThrThrGluSer------ 606

Qy        883 GGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGT 942
Db        607 --------------------SerSerAlaProAlaProThrProSerSerSerThrThr 619

Qy        943 AAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC--------------CCT 987
Db        620 GluSerSerSerAlaProValThrSerSerThrThrGluSerSerSerAlaProValPro 639

Qy        988 GTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATC---TCTGAGGGC 1044
Db        640 ThrProSerSerSerThrThrGluSerSerSerAlaProValProThrProSerSerSer 659

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
Db        660 ThrThrGluSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSer 679

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
Db        680 SerAlaProValThrSerSerThrThrGluSerSerSerAlaProValThrSerSerThr 699

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
Db        700 ThrGluSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSer 719

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
Db        720 AlaProValProThrProSerSerSerThrThrGluSerSerSerAlaProValProThr 739

Qy       1285 GGCTCCTTCAGC--------------------------AGCAGCTCCAGTTCCCAATCG 1317
Db        740 ProSerSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrGluSer 759

Qy       1318 AGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377
Db        760 SerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaProVal 779

Qy       1378 ATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCC 1437
Db        780 ProThrProSerSerSerThrThrGluSer------SerSerAlaProValProThrPro 797

Qy       1438 TCCGCTGGTGCCAAG--------------------CCCTGTGGCTCCAGCAGTGCTGGA 1476
Db        798 SerSerSerThrThrGluSerSerValAlaProValProThrProSerSerSerSer-As 817

Qy       1477 AAGATCCCCTGCCGCTCCATCCGGGATATCCT------AGCCCAAGTGAAGCCTCTGGGG 1530
Db        817 nIleThrSerSerAlaProSerSerThr-ProPheSerSerSerThrGluSerSerSerV 837

Qy       1531 CCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCC 1586
Db        837 alProVal-----------------ProThrProSerSerSerThrThrGluSer 849

RESULT 6
US-09-461-774-10
; Sequence 10, Application US/09461774
; Patent No. 6887481
; GENERAL INFORMATION:
;  APPLICANT: CHAN, Lily
;  APPLICANT: CHUNG, Maxey Ching Ming
```

```
;  APPLICANT: LIM, Renee Lay Hong
;  TITLE OF INVENTION: Bacterial-derived molecules and therapeutic and
;  TITLE OF INVENTION: diagnostic uses therefor
;  FILE REFERENCE: 1781-0180P
;  CURRENT APPLICATION NUMBER: US/09/461,774
;  CURRENT FILING DATE: 1999-12-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 386
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-461-774-10

Alignment Scores:
Pred. No.:              6.64e-10      Length:          386
Score:                  256.00        Matches:         129
Percent Similarity:     31.8%         Conservative:    32
Best Local Similarity:  25.4%         Mismatches:      159
Query Match:            8.3%          Indels:          187
DB:                     2             Gaps:            25

US-10-677-865A-2 (1-1590) x US-09-461-774-10 (1-386)

Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCA 169
Db          6 ProProAlaProProAlaProProSerProIleAsnProProValProProValProPro 25

Qy        170 CTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCT 229
Db         26 Leu------------ProAlaAlaPro---ArgThrLeuSerProProValProProAla 40

Qy        230 CCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCA 289
Db         41 Pro---ProSerProIleSerLeuAlaAlaProProLeu------ProProAspProPro 57

Qy        290 TTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATT------------ 337
Db         58 MetProProAlaIleTrpSerAlaLeuGluAlaProAsnProProValProProAlaPro 77

Qy        338 ------CCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTT 391
Db         78 ProGlyProAsnSerAlaProAlaProProMetProProThrProProLeuProProVal 97

Qy        392 CCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCA 451
Db         98 ProProGlySerGlyAla------------------------------ProArgProVal 107

Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAG 511
Db        108 ProAlaVal---------------------------------------------------- 110

Qy        512 GGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCC 571
Db        111 ------------------ProProMetProProAlaProLysArgMetProAlaLeuPro 124

Qy        572 AGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCG 631
Db        125 ProAlaProProAlaProProSerPro---ProThrSerTrpLeuAlaValProValPro 143

Qy        632 TCAGCTCCAACCAGCGTCCCTGTA---GTTCGGACATCCCCGACTCTCCCTGCAGTGGAG 688
Db        144 ProValProProValProProLeuProValLysMetProProSerProProValProPro 163

Qy        689 GGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTC 748
Db        164 PheProProAlaGluProGluThrProAsnProProAlaProPro--------------- 178

Qy        749 AGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAG 808
Db        179 ----------------------------------AlaProProLeuGluAsnSerPro 186

Qy        809 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 868

Db        187 ProProPro--------------------------ProValProProValPro------ 195

Qy        869 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 928
Db        196 ------Pro--------------------ValProProLeuThrLeu------------ 202

Qy        929 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 988
Db        202 ---------------------------------------------------------- 202

Qy        989 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 1048
Db        203 -----------------AsnProProValProProAlaProProAlaAlaAsnThrSer 216

Qy       1049 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTT---CCTCGGCCATTGCGT 1105
Db        217 AsnSerProLeuArgProProAlaProProAlaProProLeuLysProGlyPro------ 234

Qy       1106 TCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGG 1165
Db        235 ------------------------------ProAlaProProMetProProAlaProAsn 244

Qy       1166 GACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCA--------- 1216
Db        245 SerProAlaAlaProProSerProProSerProProValProValProPheProThrProPro 264

Qy       1217 ------------------------------GCAGCGCCGGTTCACCCTACCATCCCTGCG 1246
Db        265 GlyProProAlaProProGluProAsnSerSerProProAlaProProAlaProProAla 284

Qy       1247 GCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCA 1306
Db        285 AlaProLeuPro---GlyProSerProProAlaProProAlaPro------------Pro 299

Qy       1307 GTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTG 1366
Db        300 LeuProAsn-----------------------------------SerProAla------ 305

Qy       1367 GTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCC 1426
Db        306 ------------------AlaProPro------------GlyProProAlaTrp------ 313

Qy       1427 ATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCT 1486
Db        314 ------------ProGlyAlaProAspProProAlaProProLeuProTyrSerSerPro 329

Qy       1487 GCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACC 1546
Db        330 ProAlaPro-ProAlaCysProValPro------------GlyAlaProLeuAla----- 343

Qy       1547 CTGAAGTTTTCCTACCCCA 1565
Db        344 ----------ProLeuPro 346

RESULT 7
US-09-579-181-2
; Sequence 2, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
;  APPLICANT:  Chrivia, John
;  APPLICANT:  Yaciuk, Peter
;  TITLE OF INVENTION:  SNF2 Related CBP Activator Protein (SRCAP)
;  FILE REFERENCE:  16153-4247
;  CURRENT APPLICATION NUMBER: US/09/579,181
;  CURRENT FILING DATE:  2000-05-25
;  PRIOR APPLICATION NUMBER:  60/136,620
;  PRIOR FILING DATE:  1999-05-27
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2972
;   TYPE: PRT
```

```
;   ORGANISM: Human
US-09-579-181-2

Alignment Scores:
Pred. No.:              2.29e-09        Length:         2972
Score:                  251.50          Matches:        143
Percent Similarity:     40.6%           Conservative:   57
Best Local Similarity:  29.0%           Mismatches:     193
Query Match:            8.1%            Indels:         101
DB:                     2               Gaps:           25

US-10-677-865A-2 (1-1590) x US-09-579-181-2 (1-2972)

Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCA 169
              |||:::   :::|||:::      |||      |||   ||||||   :::|||   |||
Db       1151 ProAlaSerSerProMetProIleProAsn---SerSerProLeuAlaSerProValSer 1169

Qy        170 CTGGGAAGGGTGACTCCAGCGGCT--------------TCAGTAGCTACAGTGGCTCCA 214
                          |||   :::                |||||   |||      |||
Db       1170 SerThrValSerValProLeuSerSerSerLeuProIleSerValProThrThrLeuPro 1189

Qy        215 GCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTG---GTA 271
              |||      :::||||||      :::|||      :::   ||||||:::   |||
Db       1190 AlaProAlaSerAlaProLeuThrIlePro---------IleSerAlaProLeuThrVal 1206

Qy        272 GCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGG 331
              :::   :::   ||||||||
Db       1207 SerAlaSerGlyProAlaLeu------------------------------------- 1213

Qy        332 GGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTT 391
                     |||   |||||||||   ||||||   :::             |||
Db       1214 ------LeuThrSerValThrProPro---LeuAlaProValValProAlaAlaProGly 1230

Qy        392 CCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCA 451
              |||||||||        :::      :::      |||:::|||||||||
Db       1231 ProProSerLeuGlnProSerGlyAlaSerProSerAlaSerAlaLeuThr--------- 1247

Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAG 511
                    |||      |||   |||   :::   ||||||:::
Db       1248 ------LeuGlyLeuAlaThrAlaProSerLeuSerSerSerGlnThrPro--------- 1262

Qy        512 GGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCC 571
              |||      ||||||      |||   ::::::                    :::|||
Db       1263 GlyHisProLeuLeuLeuAlaProThrSerSer-----------------HisValPro 1276

Qy        572 AGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCG 631
                 |||:::                                    ||||||   |||
Db       1277 GlyLeuAsnSer--------------------------------ThrValAlaProAla 1285

Qy        632 TCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGC 691
                 :::|||      ||||||            :::         |||:::
Db       1286 CysSerProValLeuValProAlaSerAlaLeuAlaSerProPheProSerAlaProAsn 1305

Qy        692 CCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGA 751
              |||:::            :::      :::||||||   :::      |||   :::
Db       1306 ProAlaProAlaGlnAlaSerLeuLeuAlaProAlaSerSerAlaSerGlnAlaLeu--- 1324

Qy        752 GGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTC 811
                                  :::|||   |||   |||:::
Db       1325 ----------------------AlaThrProLeuAlaProMet---------AlaAla 1333

Qy        812 CCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACA 871
              |||      |||::::::|||           |||   |||      ||||||
Db       1334 ProGlnThrAlaIleLeuAla-----------ProSerProAlaProProLeuAlaPro 1349

Qy        872 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 931
                 |||::::::|||                |||||:::                   |||
Db       1350 LeuProValLeuAlaProSerProGlyAlaAlaProValLeu--------------Ala 1364

Qy        932 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 991
                       |||      :::      :::::::::|||      :::|||
Db       1365 SerSerGlnThrProValProValMetAlaProSerSerThrProGlyThrSerLeuAla 1384

Qy        992 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 1051
                    |||                |||      :::      ||||||
Db       1385 SerAlaSerProValProAlaProThrProValLeuAlaProSerSerThrGlnThrMet 1404

Qy       1052 TCTCCAGCA---ACCCCATCA---------TCCCCAGCCAGTCGGCAGCT----TCCTCG 1095
                 ||||||   ||||||         ||||||||||||   |||      :::
Db       1405 LeuProAlaProValProSerProLeuProSerProAlaSerThrGlnThrLeuAlaLeu 1424

Qy       1096 GCCATTGCGTTCCAGCCAGTG------GGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
                 |||   |||      |||         |||:::         |||            |||
Db       1425 AlaProAlaLeuAlaProThrLeuGlyGlySerSerProSerGlnThrLeuSerLeuGly 1444

Qy       1150 TCCACGGGCTCCAAGGGACCC--------------------TGCTCTCCCTCCAGTTCT 1188
                 :::         :::||||||                        :::|||:::||||||
Db       1445 ThrGlyAsnProGlnGlyProProheProThrGlnThrLeuSerLeuThrProAlaSerSer 1464

Qy       1189 CGAGTCCCCAGC---AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGC 1245
                 ||||||:::   :::   ::::::|||    ||||||      |||      |||
Db       1465 LeuValProThrProAlaGlnThrLeuSer-LeuAlaProGlyProProLeuGlyProTh 1484

Qy       1246 GGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCC 1305
                    |||      ||||||||||||      |||||||||   |||         ||
Db       1484 rGlnThrLeuSerLeuAlaProAlaProProLeuAlaProAla---SerProValGlyPr 1503

Qy       1306 AGTTCCCAATCGAGTG--------------GCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
                 |   |||                        |||   ||||||   ||||||:::|||
Db       1503 oAlaProAlaHisThrLeuThrLeuAlaProAlaSerSerSerAlaSerLeuAla-- 1522

Qy       1351 AAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC 1410
                 ||||||   :::   ||||||                |||         ||||||
Db       1523 ----ProAlaSerValGlnThrLeuThrLeuSerProAlaProValProThrLeuGlyPr 1541

Qy       1411 CCCGATGGCTCTCCCCATC---CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1467
                    |||      :::   |||   |||            ||||||   :::   ||
Db       1541 oAlaAlaAlaGlnThrLeuAlaLeuAlaProAlaSerThrGlnSerProAlaSerGlnAl 1561

Qy       1468 A---------GTGCTGGAAAGATCCCCTGCCGCTCCA 1495
                 |         |||:::      |||   |||||||||
Db       1561 aSerSerLeuValValSerAlaSerGlyAlaAlaPro 1573

RESULT 8
US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
;  APPLICANT:  Chrivia, John
;  APPLICANT:  Yaciuk, Peter
;  TITLE OF INVENTION:  SNF2 Related CBP Activator Protein (SRCAP)
;  FILE REFERENCE:  16153-4247
;  CURRENT APPLICATION NUMBER: US/09/579,181
;  CURRENT FILING DATE:  2000-05-25
;  PRIOR APPLICATION NUMBER:  60/136,620
;  PRIOR FILING DATE:  1999-05-27
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3118
;   TYPE: PRT
;   ORGANISM: Human
US-09-579-181-1

Alignment Scores:
Pred. No.:              2.31e-09        Length:         3118
Score:                  251.50          Matches:        143
Percent Similarity:     40.6%           Conservative:   57
Best Local Similarity:  29.0%           Mismatches:     193
Query Match:            8.1%            Indels:         101
DB:                     2               Gaps:           25

US-10-677-865A-2 (1-1590) x US-09-579-181-1 (1-3118)

Qy        110 CCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCA 169
              |||:::   :::|||:::      |||      |||   ||||||   :::|||   |||
```

```
Db       1297 ProAlaSerSerProMetProIleProAsn---SerSerProLeuAlaSerProValSer 1315

Qy        170 CTGGGAAGGGTGACTCCAGCGGCT--------------TCAGTAGCTACAGTGGCTCCA 214

Db       1316 SerThrValSerValProLeuSerSerSerLeuProIleSerValProThrThrLeuPro 1335

Qy        215 GCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTG---GTA 271

Db       1336 AlaProAlaSerAlaProLeuThrIlePro---------IleSerAlaProLeuThrVal 1352

Qy        272 GCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGG 331

Db       1353 SerAlaSerGlyProAlaLeu------------------------------------- 1359

Qy        332 GGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTT 391

Db       1360 ------LeuThrSerValThrProPro---LeuAlaProValValProAlaAlaProGly 1376

Qy        392 CCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCA 451

Db       1377 ProProSerLeuGlnProSerGlyAlaSerProSerAlaSerAlaLeuThr--------- 1393

Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAG 511

Db       1394 ------LeuGlyLeuAlaThrAlaProSerLeuSerSerSerGlnThrPro--------- 1408

Qy        512 GGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCTTCCC 571

Db       1409 GlyHisProLeuLeuLeuAlaProThrSerSer------------------HisValPro 1422

Qy        572 AGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCG 631

Db       1423 GlyLeuAsnSer--------------------------------ThrValAlaProAla 1431

Qy        632 TCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGC 691

Db       1432 CysSerProValLeuValProAlaSerAlaLeuAlaSerProPheProSerAlaProAsn 1451

Qy        692 CCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGA 751

Db       1452 ProAlaProAlaGlnAlaSerLeuLeuAlaProAlaSerSerAlaSerGlnAlaLeu--- 1470

Qy        752 GGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTC 811

Db       1471 ------------------------AlaThrProLeuAlaProMet---------AlaAla 1479

Qy        812 CCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACA 871

Db       1480 ProGlnThrAlaIleLeuAla------------ProSerProAlaProProLeuAlaPro 1495

Qy        872 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 931

Db       1496 LeuProValLeuAlaProSerProGlyAlaAlaProValLeu---------------Ala 1510

Qy        932 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 991

Db       1511 SerSerGlnThrProValProValMetAlaProSerSerThrProGlyThrSerLeuAla 1530

Qy        992 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 1051

Db       1531 SerAlaSerProValProAlaProThrProValLeuAlaProSerSerThrGlnThrMet 1550

Qy       1052 TCTCCAGCA---ACCCCATCA---------TCCCCAGCCAGTCGGCAGCT----TCCTCG 1095

Db       1551 LeuProAlaProValProSerProLeuProSerProAlaSerThrGlnThrLeuAlaLeu 1570

Qy       1096 GCCATTGCGTTCCAGCCAGTG------GGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149

Db       1571 AlaProAlaLeuAlaProThrLeuGlyGlySerSerProSerGlnThrLeuSerLeuGly 1590

Qy       1150 TCCACGGGCTCCAAGGGACCC--------------------TGCTCTCCCTCCAGTTCT 1188

Db       1591 ThrGlyAsnProGlnGlyProPheProThrGlnThrLeuSerLeuThrProAlaSerSer 1610

Qy       1189 CGAGTCCCCAGC---AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGC 1245

Db       1611 LeuValProThrProAlaGlnThrLeuSer-LeuAlaProGlyProProLeuGlyProTh 1630

Qy       1246 GGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCC 1305

Db       1630 rGlnThrLeuSerLeuAlaProAlaProProLeuAlaProAla---SerProValGlyPr 1649

Qy       1306 AGTTCCCAATCGAGTG--------------GCAAAATCATCCTTCAGCCTTGTGGCAGC 1350

Db       1649 oAlaProAlaHisThrLeuThrLeuAlaProAlaSerSerSerAlaSerLeuLeuAla-- 1668

Qy       1351 AAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC 1410

Db       1669 ----ProAlaSerValGlnThrLeuThrLeuSerProAlaProValProThrLeuGlyPr 1687

Qy       1411 CCCGATGGCTCTCCCCATC---CTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1467

Db       1687 oAlaAlaAlaGlnThrLeuAlaLeuAlaProAlaSerThrGlnSerProAlaSerGlnAl 1707

Qy       1468 A---------GTGCTGGAAAGATCCCCTGCCGCTCCA 1495

Db       1707 aSerSerLeuValValSerAlaSerGlyAlaAlaPro 1719
```

RESULT 9
US-08-556-978B-59
; Sequence 59, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FAHNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; TITLE OF INVENTION: SPIDER SILK ANALOGS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-556-978B-59

Alignment Scores:			
Pred. No.:	2.76e-09	Length:	493
Score:	247.50	Matches:	121
Percent Similarity:	36.4%	Conservative:	51

```
Best Local Similarity:  25.6%          Mismatches:    216
Query Match:            8.0%           Indels:        85
DB:                     2              Gaps:          20

US-10-677-865A-2 (1-1590) x US-08-556-978B-59 (1-493)

Qy        172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231
Db         62 GlyTyrGlyGlnGlyGlnGlnGlyProSerGlyProGlySerAlaAlaAlaAlaSerAla 81

Qy        232 ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291
Db         82 AlaAlaSerAlaGluSerGlyGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProGly 101

Qy        292 GCCCAGGGT----------------------------------GGTTCTGCAGGATCTTTT 318
Db        102 GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSer 121

Qy        319 AAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAA 378
Db        122 GlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSerGlyProGlyGlnGlnGlyPro 141

Qy        379 GGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCT 438
Db        142 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 161

Qy        439 CACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGC 498
Db        162 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSer------------------ 175

Qy        499 AGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATA 558
Db        176 -----------GlyProGlyGlnGlnGlyPro--------------------GlyGly 184

Qy        559 CTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGC 618
Db        185 TyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyThrSerGly 204

Qy        619 AGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCC 678
Db        205 ProGlySerAlaAlaAlaAla---------------------------------------- 211

Qy        679 TGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTG 738
Db        212 AlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 231

Qy        739 TCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT------GCC 792
Db        232 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlyTyrGly 251

Qy        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
Db        252 ProGly---GlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSerGlyPro 270

Qy        853 CCAATCACCTCTGTAGACAAATCCTATGGT---GGCTACGAGGTGGTGGGTGGCTCCTCT 909
Db        271 GlySerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlnGlnGlyLeuGlyGlyTyrGly 290

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
Db        291 ProGlyGlnGlnGlyProGly---GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyr 309

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
Db        310 -----------GlyPro------GlySerAlaSerAlaAlaAlaAlaAlaAlaGlyPro 323

Qy       1030 ------------------CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATC 1071
Db        324 GlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSerGlyPro------ 341

Qy       1072 CCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGT---GGG 1128
Db        342 GlySerAlaSerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlyTyrGlyProGly 361

Qy       1129 GTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCT 1188
Db        362 GlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSerGlyProGlySer--- 380

Qy       1189 CGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCC------TACCATCCC 1242
Db        381 ---------AlaSerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlyTyrGlyPro 397

Qy       1243 ------------TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCC 1290
Db        398 GlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSerGlyProGlySer 417

Qy       1291 TTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
Db        418 ---AlaAlaAlaAlaAlaAlaSerAlaGly---------------ProGlyGlyTyr 431

Qy       1351 AAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGC 1410
Db        432 GlyProAlaGlnGlnGlyProSerGlyProGlyIleAlaAlaSerAlaAlaSerAlaGly 451

Qy       1411 CCCGATGGC---TCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCC------------ 1455
Db        452 ProGlyGlyTyrGlyProAlaGlnGlnGlyProAlaGlyTyrGlyProGlySerAlaVal 471

Qy       1456 ---------TGTGGCTCCAGCAGTGCTGGAAAGATCCCC 1485
Db        472 AlaAlaSerAlaGlyAlaGlySerAlaGlyTyrGlyPro 484

RESULT 10
US-09-581-472B-2
; Sequence 2, Application US/09581472B
; Patent No. 6818744
; GENERAL INFORMATION:
;  APPLICANT: FRADE, Raymond
;  TITLE OF INVENTION: p53 REGULATORY PROTEIN CALLED RB18A AND USES THEREOF
;  FILE REFERENCE: P06781US00/BAS
;  CURRENT APPLICATION NUMBER: US/09/581,472B
;  CURRENT FILING DATE:  2000-08-14
;  PRIOR APPLICATION NUMBER: PCT/EP98/08560
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: EP 97403051.2
;  PRIOR FILING DATE: 1997-12-15
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 1566
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-581-472B-2

Alignment Scores:
Pred. No.:              3.69e-09       Length:        1566
Score:                  247.50         Matches:       138
Percent Similarity:     34.8%          Conservative:  34
Best Local Similarity:  27.9%          Mismatches:    173
Query Match:            8.0%           Indels:        149
DB:                     2              Gaps:          22

US-10-677-865A-2 (1-1590) x US-09-581-472B-2 (1-1566)

Qy         40 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 99
Db        960 GlyAsnGlyThrSerAsnSerThrLeuSerGly---------ProGlyLeuAspSerLys 976

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
Db        977 ProGlyLysArgSerArgThrProSerAsnAspGly---------LysSerLysAspLys 993

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
Db        994 ProProLysArgLysLysAlaAspThrGluGlyLysSer------ProSerHisSerSer 1011
```

```
Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
Db       1012 SerAsnArgProPheThrProProThrSerThrGlyGlySerLysSerProGlySerAla 1031

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
Db       1032 GlyArgSerGlnThrProProGlyValAlaThrProProIleProLysIle--------- 1048

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
Db       1049 ----------------------------ThrIleGlnIleProLysGlyThrValMet 1058

Qy        400 CTGGGGAGCAGCAGCTCTCACTCG------GGAAGCAGCGGC------TCTCACTCGGGA 447
Db       1059 ValGlyLysProSerSerHisSerGlnTyrThrSerSerGlySerValSerSerSerGly 1078

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507
Db       1079 SerLysSerHisHisSerHisSerSerSerSerSerSerSerAlaSerThrSer------ 1096

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567
Db       1097 ---------------------------------------------GlyLysMetLysSer 1101

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627
Db       1102 SerLysSerGluGlySerSerSerSerLysLeuSerSerSerMetTyrSerSerGlnGly 1121

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687
Db       1122 SerSerGlySerSerGlnSerLysAsnSerSerGlnSerGlyGlyLysProGlySer--- 1140

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGT 747
Db       1141 SerProIleThrLysHisGly------------LeuSerSerGlySerSerSerThrLys 1156

Qy        748 CAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAA 807
Db       1157 MetLysPro--------------------------GlnGlyLysProSerSerLeuMet 1167

Qy        808 GGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAGCCCTGTCCCCCAATCACCTCT 864
Db       1168 AsnProSerLeuSerLysProAsnIleSerProSerHisSerArgProPro--------- 1184

Qy        865 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 924
Db       1185 ----------------------------------GlyGlySerAspLysLeuAlaSer 1192

Qy        925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984
Db       1193 Pro-------------------------------------------------------MetLys 1195

Qy        985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
Db       1196 ProValProGlyThrPro------ProSerSerLysAlaLysSerProIleSerSer--- 1212

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
Db       1212 ------------------------------------------------------------ 1212

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
Db       1213 -----------GlySerGlyGlySerHisMetSerGly--------------------- 1221

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
Db       1222 ---------ThrSerSerSerSerGlyMetLysSerSerSerGlyLeuGlySerSer--- 1237

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
Db       1238 --------------------GlySerLeuSerGlnLys------ThrProProSerSer 1248

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCC---CAATCGAGTGGCAAAATCATCCTTCAGCCT 1341
```

```
Db       1249 AsnSerCysThrAlaSerSerSerSerPheSerSerSerGlySerSerMetSerSerSer 1268

Qy       1342 TGTGGCAGCAAGTCCAGCTCTTCTGGT---CACCCTTGCATGTCTGTCTCC---TCCTTG 1395
Db       1269 GlnAsnGlnHisGlySerSerLysGlyLysSerProSerArgAsnLysLysProSerLeu 1288

Qy       1396 ACA----------------------------CTGACTGGGGGCCCCGATGGCTCTCCC 1425
Db       1289 ThrAlaValIleAspLysLeuLysHisGlyValValThrSerGlyProGlyGlyGluAsp 1308

Qy       1426 CATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1467
Db       1309 ProLeuAspGlyGlnMetGlyValSerThrAsnSerSerSer 1322

RESULT 11
US-10-110-517-2
; Sequence 2, Application US/09110517A
; Patent No. 6248520
; GENERAL INFORMATION:
;  APPLICANT: Roeder, Robert G
;  APPLICANT: Fondell, Joseph D
;  APPLICANT: Yuan, Chao X
;  APPLICANT: Ito, Mitsuhiro
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NUCLEAR HORMONE
;  TITLE OF INVENTION: RECEPTOR COACTIVATORS AND USES THEREOF
;  FILE REFERENCE: 600-1-224
;  CURRENT APPLICATION NUMBER: US/09/110,517A
;  CURRENT FILING DATE: 1998-07-06
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1581
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-110-517-2

Alignment Scores:
Pred. No.:              3.7e-09      Length:              1581
Score:                  247.50       Matches:             138
Percent Similarity:     34.8%        Conservative:        34
Best Local Similarity:  27.9%        Mismatches:          173
Query Match:            8.0%         Indels:              149
DB:                     2            Gaps:                22

US-10-677-865A-2 (1-1590) x US-09-110-517-2 (1-1581)

Qy         40 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 99
Db        975 GlyAsnGlyThrSerAsnSerThrLeuSerGly---------ProGlyLeuAspSerLys 991

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
Db        992 ProGlyLysArgSerArgThrProSerAsnAspGly---------LysSerLysAspLys 1008

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
Db       1009 ProProLysArgLysLysAlaAspThrGluGlyLysSer------ProSerHisSerSer 1026

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
Db       1027 SerAsnArgProPheThrProProThrSerThrGlyGlySerLysSerProGlySerAla 1046

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
Db       1047 GlyArgSerGlnThrProProGlyValAlaThrProProIleProLysIle--------- 1063

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
Db       1064 ----------------------------ThrIleGlnIleProLysGlyThrValMet 1073

Qy        400 CTGGGGAGCAGCAGCTCTCACTCG------GGAAGCAGCGGC------TCTCACTCGGGA 447
```

```
Db       1074 ValGlyLysProSerSerHisSerGlnTyrThrSerSerGlySerValSerSerSerGly 1093

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507

Db       1094 SerLysSerHisHisSerHisSerSerSerSerSerSerSerAlaSerThrSer------ 1111

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567

Db       1112 ----------------------------------------------GlyLysMetLysSer 1116

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627

Db       1117 SerLysSerGluGlySerSerSerSerLysLeuSerSerSerMetTyrSerSerGlnGly 1136

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687

Db       1137 SerSerGlySerSerGlnSerLysAsnSerSerGlnSerGlyGlyLysProGlySer--- 1155

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGT 747

Db       1156 SerProIleThrLysHisGly------------LeuSerSerGlySerSerSerThrLys 1171

Qy        748 CAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAA 807

Db       1172 MetLysPro--------------------------GlnGlyLysProSerSerLeuMet 1182

Qy        808 GGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAGCCCTGTCCCCCAATCACCTCT 864

Db       1183 AsnProSerLeuSerLysProAsnIleSerProSerHisSerArgProPro--------- 1199

Qy        865 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 924

Db       1200 ----------------------------------GlyGlySerAspLysLeuAlaSer 1207

Qy        925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984

Db       1208 Pro-------------------------------------------------MetLys 1210

Qy        985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044

Db       1211 ProValProGlyThrPro------ProSerSerLysAlaLysSerProIleSerSer--- 1227

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104

Db       1227 ------------------------------------------------------------ 1227

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164

Db       1228 -----------GlySerGlyGlySerHisMetSerGly-------------------- 1236

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224

Db       1237 ---------ThrSerSerSerSerGlyMetLysSerSerSerGlyLeuGlySerSer--- 1252

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284

Db       1253 --------------------GlySerLeuSerGlnLys------ThrProProSerSer 1263

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCC---CAATCGAGTGGCAAAATCATCCTTCAGCCT 1341

Db       1264 AsnSerCysThrAlaSerSerSerSerPheSerSerSerGlySerSerMetSerSerSer 1283

Qy       1342 TGTGGCAGCAAGTCCAGCTCTTCTGGT---CACCCTTGCATGTCTGTCTCC---TCCTTG 1395

Db       1284 GlnAsnGlnHisGlySerSerLysGlyLysSerProSerArgAsnLysLysProSerLeu 1303

Qy       1396 ACA----------------------------CTGACTGGGGGCCCCGATGGCTCTCCC 1425

Db       1304 ThrAlaValIleAspLysLeuLysHisGlyValValThrSerGlyProGlyGlyGluAsp 1323

Qy       1426 CATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1467

Db       1324 ProLeuAspGlyGlnMetGlyValSerThrAsnSerSerSer 1337
```

```
RESULT 12
US-08-425-069-4
; Sequence 4, Application US/08425069
; Patent No. 5728810
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Xu, Ming
;    APPLICANT:  Hinman, Michael B.
;    TITLE OF INVENTION:  ISOLATED DNA CODING FOR SPIDER SILK
;    TITLE OF INVENTION:  PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
;    TITLE OF INVENTION:  CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  301 No. 5728810th Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22046
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/425,069
;      FILING DATE:  19-APR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-106P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  595 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-425-069-4

Alignment Scores:
Pred. No.:                4.32e-09        Length:          595
Score:                    245.00          Matches:         126
Percent Similarity:       36.3%           Conservative:    56
Best Local Similarity:    25.1%           Mismatches:      217
Query Match:              7.9%            Indels:          102
DB:                       1               Gaps:            23

US-10-677-865A-2 (1-1590) x US-08-425-069-4 (1-595)

Qy        172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231

Db         98 GlyTyrGlyGlnGlyGlnGlnGlyProSerGlyProGlySerAlaAlaAlaAlaSerAla 117

Qy        232 ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291

Db        118 AlaAlaSerAlaGluSerGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 137

Qy        292 GCCCAGGGT-------------------------------------GGTTCTGCAGGA 312

Db        138 ProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 157

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372

Db        158 ProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSerGlyProGlyGlnGln 177

Qy        373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGC 432
```

```
Db        178 GlyProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGln 197

Qy        433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492

Db        198 GlyProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSer------------ 213

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552

Db        214 -----------------GlyProGlyGlnGlnGlyPro--------------------- 220

Qy        553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612

Db        221 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyLeu 240

Qy        613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672

Db        241 SerGlyProGlySerAlaAlaAlaAla------------------------------- 249

Qy        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732

Db        250 ------AlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGln 267

Qy        733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT--- 789

Db        268 GlyProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlyGly 287

Qy        790 ---GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAG 843

Db        288 TyrGlyProGly---GlnGlnGlyPro-----------GlyGlyTyrGlyProGlyGln 302

Qy        844 CCCTGTCCC-----------CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891

Db        303 GlnGlyProSerGlyAlaGlySerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlnGln 322

Qy        892 GTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAA 951

Db        323 GlyLeuGlyGlyTyrGlyProGlyGlnGlnGlyProGly---GlyTyrGlyProGlyGln 341

Qy        952 ATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCT 1011

Db        342 GlnGlyProGlyGlyTyr-----------GlyPro------GlySerAlaSerAlaAla 355

Qy       1012 TCCTTTGCAGCTGGGCCC-----------------CCCATCTCTGAGGGCAAATACTTC 1053

Db        356 AlaAlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 375

Qy       1054 TCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCA 1113

Db        376 ProSerGlyPro------GlySerAlaSerAlaAlaAlaAlaAlaAlaAlaAlaGlyPro 393

Qy       1114 GTGGGGACTGGT---GGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCC 1170

Db        394 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyPro 413

Qy       1171 TGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCA 1230

Db        414 SerGlyProGlySer-----------AlaSerAlaAlaAlaAlaAlaAlaAlaAlaGly 429

Qy       1231 CCC------TACCATCCC-----------TGCGGCAGTGCTTCCCAGAGCCCCTGCTCC 1272

Db        430 ProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGly 449

Qy       1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATC 1332

Db        450 ProSerGlyProGlySer---AlaAlaAlaAlaAlaAlaAlaSerAlaGly--------- 465

Qy       1333 CTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCC 1392

Db        466 ------ProGlyGlyTyrGlyProAlaGlnGlnGlyProSerGlyProGlyIleAlaAla 483

Qy       1393 TTGACACTGACTGGGGGCCCCGATGGC---TCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449

Db        484 SerAlaAlaSerAlaGlyProGlyGlyTyrGlyProAlaGlnGlnGlyProAlaGlyTyr 503

Qy       1450 AAGCCC--------------------TGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488

Db        504 GlyProGlySerAlaValAlaAlaSerAlaGlyAlaGlySerAlaGlyTyrGlyProGly 523

Qy       1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCT 1548

Db        524 --------------------SerGlnAlaSerAlaAlaAlaSerArgLeuAlaSerPro 536

Qy       1549 GAA 1551

Db        537 Asp 537
```

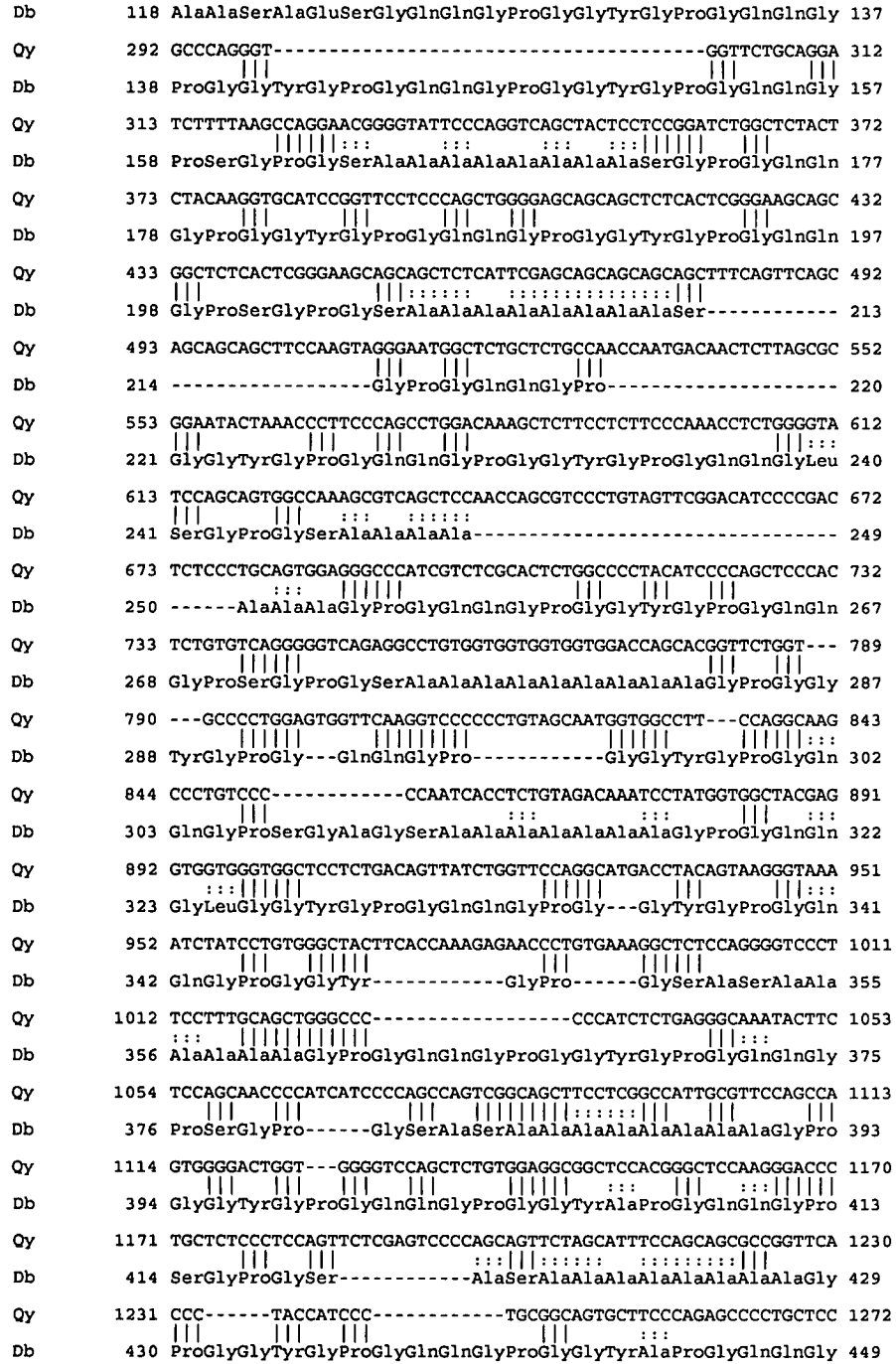
RESULT 13
US-08-317-844B-4
; Sequence 4, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-317-844B-4

Alignment Scores:

Pred. No.:	4.32e-09	Length:	595
Score:	245.00	Matches:	126
Percent Similarity:	36.3%	Conservative:	56
Best Local Similarity:	25.1%	Mismatches:	217
Query Match:	7.9%	Indels:	102
DB:	1	Gaps:	23

US-10-677-865A-2 (1-1590) x US-08-317-844B-4 (1-595)

```
Qy        172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231

Db         98 GlyTyrGlyGlnGlyGlnGlnGlyProSerGlyProGlySerAlaAlaAlaAlaSerAla 117

Qy        232 ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291
```


Db 118 AlaAlaSerAlaGluSerGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 137
Qy 292 GCCCAGGGT---------- GGTTCTGCAGGA 312
Db 138 ProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 157
Qy 313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372
Db 158 ProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSerGlyProGlyGlnGln 177
Qy 373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGC 432
Db 178 GlyProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGln 197
Qy 433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492
Db 198 GlyProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSer---------- 213
Qy 493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552
Db 214 ----------GlyProGlyGlnGlnGlyPro---------- 220
Qy 553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612
Db 221 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyLeu 240
Qy 613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672
Db 241 SerGlyProGlySerAlaAlaAlaAla---------- 249
Qy 673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
Db 250 ------AlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGln 267
Qy 733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT--- 789
Db 268 GlyProSerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlyGly 287
Qy 790 ---GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAG 843
Db 288 TyrGlyProGly---GlnGlnGlyPro----------GlyGlyTyrGlyProGlyGln 302
Qy 844 CCCTGTCCC---------- CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891
Db 303 GlnGlyProSerGlyAlaGlySerAlaAlaAlaAlaAlaAlaGlyProGlyGlnGln 322
Qy 892 GTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAA 951
Db 323 GlyLeuGlyGlyTyrGlyProGlyGlnGlnGlyProGly---GlyTyrGlyProGlyGln 341
Qy 952 ATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCT 1011
Db 342 GlnGlyProGlyGlyTyr----------GlyPro------GlySerAlaSerAlaAla 355
Qy 1012 TCCTTTGCAGCTGGGCCC---------- CCCATCTCTGAGGGCAAATACTTC 1053
Db 356 AlaAlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGly 375
Qy 1054 TCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCA 1113
Db 376 ProSerGlyPro------GlySerAlaSerAlaAlaAlaAlaAlaAlaAlaGlyPro 393
Qy 1114 GTGGGGACTGGT---GGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCC 1170
Db 394 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyPro 413
Qy 1171 TGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCA 1230
Db 414 SerGlyProGlySer----------AlaSerAlaAlaAlaAlaAlaAlaAlaAlaGly 429
Qy 1231 CCC------TACCATCCC---------- TGCGGCAGTGCTTCCCAGAGCCCCTGCTCC 1272
Db 430 ProGlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGly 449
Qy 1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATC 1332
Db 450 ProSerGlyProGlySer---AlaAlaAlaAlaAlaAlaAlaSerAlaGly--------- 465
Qy 1333 CTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCC 1392
Db 466 ------ProGlyGlyTyrGlyProAlaGlnGlnGlyProSerGlyProGlyIleAlaAla 483
Qy 1393 TTGACACTGACTGGGGGCCCCGATGGC---TCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
Db 484 SerAlaAlaSerAlaGlyProGlyGlyTyrGlyProAlaGlnGlnGlyProAlaGlyTyr 503
Qy 1450 AAGCCC---------- TGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGC 1488
Db 504 GlyProGlySerAlaValAlaAlaSerAlaGlyAlaGlySerAlaGlyTyrGlyProGly 523
Qy 1489 CGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCT 1548
Db 524 ----------SerGlnAlaSerAlaAlaAlaSerArgLeuAlaSerPro 536
Qy 1549 GAA 1551
Db 537 Asp 537

RESULT 14
US-09-247-806-2
; Sequence 2, Application US/09247806
; Patent No. 6280747
; GENERAL INFORMATION:
; APPLICANT: PHILLIPPE, Michel
; APPLICANT: GARSON, Jean-Claude
; APPLICANT: ARRAUDEAU, Jean-Pierre
; TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
; TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
; TITLE OF INVENTION: ANALOG
; FILE REFERENCE: 6388-0365-0
; CURRENT APPLICATION NUMBER: US/09/247,806
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: FR 98/01614
; EARLIER FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-247-806-2

Alignment Scores:

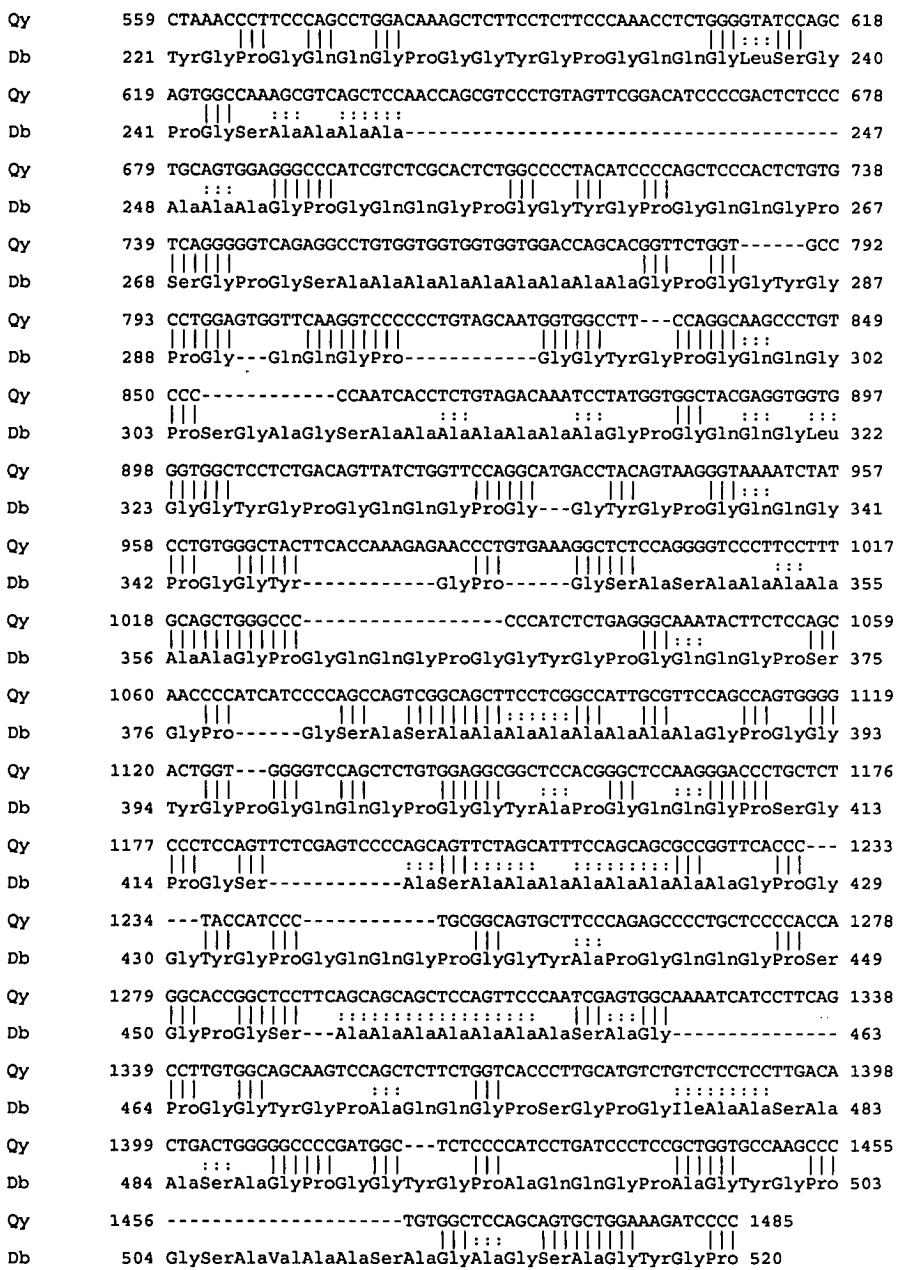
Pred. No.:	5.34e-09	Length:	529
Score:	243.50	Matches:	122
Percent Similarity:	36.7%	Conservative:	53
Best Local Similarity:	25.6%	Mismatches:	209
Query Match:	7.8%	Indels:	93
DB:	2	Gaps:	22

US-10-677-865A-2 (1-1590) x US-09-247-806-2 (1-529)

Qy 172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231
Db 98 GlyTyrGlyGlnGlyGlnGlnGlyProSerGlyProGlySerAlaAlaAlaAlaSerAla 117
Qy 232 ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291
Db 118 AlaAlaSerAlaGluSerGlyGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProGly 137
Qy 292 GCCCAGGGT---------- GGTTCTGCAGGATCTTTT 318
Db 138 GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSer 157
Qy 319 AAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAA 378

```
Db        158 GlyProGlySerAlaAlaAlaAlaAlaAlaAlaSerGlyProGlyGlnGlnGlyPro 177

Qy        379 GGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCT 438

Db        178 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 197

Qy        439 CACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGC 498

Db        198 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSer------------------ 211

Qy        499 AGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATA 558

Db        212 -----------GlyProGlyGlnGlnGlyPro--------------------GlyGly 220

Qy        559 CTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGC 618

Db        221 TyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyLeuSerGly 240

Qy        619 AGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCC 678

Db        241 ProGlySerAlaAlaAlaAla---------------------------------------- 247

Qy        679 TGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTG 738

Db        248 AlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 267

Qy        739 TCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT------GCC 792

Db        268 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlyTyrGly 287

Qy        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAGCCCTGT 849

Db        288 ProGly---GlnGlnGlyPro-----------GlyGlyTyrGlyProGlyGlnGlnGly 302

Qy        850 CCC-----------CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTG 897

Db        303 ProSerGlyAlaGlySerAlaAlaAlaAlaAlaAlaGlyProGlyGlnGlnGlyLeu 322

Qy        898 GGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAT 957

Db        323 GlyGlyTyrGlyProGlyGlnGlnGlyProGly---GlyTyrGlyProGlyGlnGlnGly 341

Qy        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTT 1017

Db        342 ProGlyGlyTyr-----------GlyPro------GlySerAlaSerAlaAlaAlaAla 355

Qy       1018 GCAGCTGGGCCC-----------------CCCATCTCTGAGGGCAAATACTTCTCCAGC 1059

Db        356 AlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSer 375

Qy       1060 AACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGG 1119

Db        376 GlyPro------GlySerAlaSerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGly 393

Qy       1120 ACTGGT---GGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCT 1176

Db        394 TyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSerGly 413

Qy       1177 CCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCC--- 1233

Db        414 ProGlySer-----------AlaSerAlaAlaAlaAlaAlaAlaAlaAlaGlyProGly 429

Qy       1234 ---TACCATCCC-----------TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCA 1278

Db        430 GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSer 449

Qy       1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAG 1338

Db        450 GlyProGlySer---AlaAlaAlaAlaAlaAlaAlaSerAlaGly-------------- 463

Qy       1339 CCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACA 1398

Db        464 ProGlyGlyTyrGlyProAlaGlnGlnGlyProSerGlyProGlyIleAlaAlaSerAla 483
```

```
Qy       1399 CTGACTGGGGGCCCCGATGGC---TCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCC 1455

Db        484 AlaSerAlaGlyProGlyGlyTyrGlyProAlaGlnGlnGlyProAlaGlyTyrGlyPro 503

Qy       1456 --------------------TGTGGCTCCAGCAGTGCTGGAAAGATCCCC 1485

Db        504 GlySerAlaValAlaAlaSerAlaGlyAlaGlySerAlaGlyTyrGlyPro 520

RESULT 15
US-09-861-597-2
; Sequence 2, Application US/09861597
; Patent No. 6841162
; GENERAL INFORMATION:
;  APPLICANT: PHILLIPPE, Michel
;  APPLICANT:  GARSON, Jean-Claude
;  APPLICANT:  ARRAUDEAU, Jean-Pierre
;  TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
;  TITLE OF INVENTION:  LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
;  TITLE OF INVENTION:  ANALOG
;  FILE REFERENCE: 6388-0365-0
;  CURRENT APPLICATION NUMBER: US/09/861,597
;  CURRENT FILING DATE:  2001-05-22
;  PRIOR APPLICATION NUMBER: 09/247,806
;  PRIOR FILING DATE: 1999-02-11
;  PRIOR APPLICATION NUMBER: FR 98/01614
;  PRIOR FILING DATE: 1998-02-11
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Nephila clavipes
US-09-861-597-2

Alignment Scores:
Pred. No.:                5.34e-09      Length:          529
Score:                    243.50        Matches:         122
Percent Similarity:       36.7%         Conservative:    53
Best Local Similarity:    25.6%         Mismatches:      209
Query Match:              7.8%          Indels:          93
DB:                       2             Gaps:            22

US-10-677-865A-2 (1-1590) x US-09-861-597-2 (1-529)

Qy        172 GGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCC 231

Db         98 GlyTyrGlyGlnGlyGlnGlnGlyProSerGlyProGlySerAlaAlaAlaAlaSerAla 117

Qy        232 ATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291

Db        118 AlaAlaSerAlaGluSerGlyGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProGly 137

Qy        292 GCCCAGGGT---------------------------------GGTTCTGCAGGATCTTTT 318

Db        138 GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSer 157

Qy        319 AAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAA 378

Db        158 GlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSerGlyProGlyGlnGlnGlyPro 177

Qy        379 GGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCT 438

Db        178 GlyGlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 197

Qy        439 CACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGC 498

Db        198 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaSer------------------ 211

Qy        499 AGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATA 558

Db        212 -----------GlyProGlyGlnGlnGlyPro--------------------GlyGly 220
```


Qy 559 CTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGC 618
Db 221 TyrGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyLeuSerGly 240
Qy 619 AGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCC 678
Db 241 ProGlySerAlaAlaAlaAla---------- 247
Qy 679 TGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTG 738
Db 248 AlaAlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyPro 267
Qy 739 TCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT------GCC 792
Db 268 SerGlyProGlySerAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlyTyrGly 287
Qy 793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT---CCAGGCAAGCCCTGT 849
Db 288 ProGly---GlnGlnGlyPro----------GlyGlyTyrGlyProGlyGlnGlnGly 302
Qy 850 CCC----------CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTG 897
Db 303 ProSerGlyAlaGlySerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGlnGlnGlyLeu 322
Qy 898 GGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAT 957
Db 323 GlyGlyTyrGlyProGlyGlnGlnGlyProGly---GlyTyrGlyProGlyGlnGlnGly 341
Qy 958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTT 1017
Db 342 ProGlyGlyTyr----------GlyPro------GlySerAlaSerAlaAlaAlaAla 355
Qy 1018 GCAGCTGGGCCC----------CCCATCTCTGAGGGCAAATACTTCTCCAGC 1059
Db 356 AlaAlaGlyProGlyGlnGlnGlyProGlyGlyTyrGlyProGlyGlnGlnGlyProSer 375
Qy 1060 AACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGG 1119
Db 376 GlyPro------GlySerAlaSerAlaAlaAlaAlaAlaAlaAlaGlyProGlyGly 393
Qy 1120 ACTGGT---GGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCT 1176
Db 394 TyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSerGly 413
Qy 1177 CCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCC--- 1233
Db 414 ProGlySer----------AlaSerAlaAlaAlaAlaAlaAlaAlaAlaGlyProGly 429
Qy 1234 ---TACCATCCC----------TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCA 1278
Db 430 GlyTyrGlyProGlyGlnGlnGlyProGlyGlyTyrAlaProGlyGlnGlnGlyProSer 449
Qy 1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAG 1338
Db 450 GlyProGlySer---AlaAlaAlaAlaAlaAlaAlaSerAlaGly---------- 463
Qy 1339 CCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACA 1398
Db 464 ProGlyGlyTyrGlyProAlaGlnGlnGlyProSerGlyProGlyIleAlaAlaSerAla 483
Qy 1399 CTGACTGGGGGCCCCGATGGC---TCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCC 1455
Db 484 AlaSerAlaGlyProGlyGlyTyrGlyProAlaGlnGlnGlyProAlaGlyTyrGlyPro 503
Qy 1456 ----------TGTGGCTCCAGCAGTGCTGGAAAGATCCCC 1485
Db 504 GlySerAlaValAlaAlaSerAlaGlyAlaGlySerAlaGlyTyrGlyPro 520

Search completed: March 21, 2006, 22:09:26
Job time : 60.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:58 ; Search time 47.4182 Seconds
(without alignments)
238.062 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	92.9	529	1	CDSN_HUMAN	Q15517 homo sapien
2	78	92.9	529	1	CDSN_PANTR	Q7yr44 pan troglod
3	78	92.9	534	1	CDSN_MACMU	Q5tm45 macaca mula
4	65	77.4	561	1	CDSN_MOUSE	Q7tpc1 mus musculu
5	59	70.2	225	1	CDSN_PIG	O19084 sus scrofa
6	49	58.3	293	2	Q6NUU9_BRARE	Q6nuu9 brachydanio
7	48	57.1	283	1	AROE_METTH	O26344 methanobact
8	48	57.1	367	2	Q6YXZ0_ORYSA	Q6yxz0 oryza sativ
9	47	56.0	1473	2	Q4HZJ7_GIBZE	Q4hzj7 gibberella
10	46	54.8	115	2	Q9DDJ7_CHICK	Q9ddj7 gallus gall
11	46	54.8	289	1	TXPH_PYETR	Q07338 pyemotes tr
12	46	54.8	387	2	Q857A1_9CAUD	Q857a1 mycobacteri
13	46	54.8	540	2	Q8XLS8_CLOPE	Q8xls8 clostridium
14	46	54.8	682	2	Q7VF32_HELHP	Q7vf32 helicobacte
15	45	53.6	138	2	Q9I3U9_PSEAE	Q9i3u9 pseudomonas
16	45	53.6	197	2	Q9XC66_STRFR	Q9xc66 streptomyce
17	45	53.6	259	2	Q86JY1_DICDI	Q86jy1 dictyosteli
18	45	53.6	279	1	TXP1_PYETR	P20798 pyemotes tr
19	45	53.6	317	2	Q4K2A1_STRPN	Q4k2a1 streptococc
20	45	53.6	320	2	Q4K2C7_STRPN	Q4k2c7 streptococc
21	45	53.6	413	1	GATA1_MOUSE	P17679 mus musculu
22	45	53.6	413	2	Q7TMX8_MOUSE	Q7tmx8 mus musculu
23	45	53.6	423	2	Q7SAX8_NEUCR	Q7sax8 neurospora
24	45	53.6	426	2	Q5QQ59_PSEFL	Q5qq59 pseudomonas
25	45	53.6	458	2	Q7S663_NEUCR	Q7s663 neurospora
26	45	53.6	537	2	O87201_CORGL	O87201 corynebacte
27	45	53.6	537	2	Q8NSG8_CORGL	Q8nsg8 corynebacte
28	45	53.6	537	2	Q8RQN2_COREF	Q8rqn2 corynebacte
29	45	53.6	543	2	Q6NIV8_CORDI	Q6niv8 corynebacte
30	45	53.6	642	2	Q7NQ35_CHRVO	Q7nq35 chromobacte
31	44	52.4	93	2	Q6ET27_ORYSA	Q6et27 oryza sativ
32	44	52.4	126	2	Q7Z761_HUMAN	Q7z761 homo sapien
33	44	52.4	267	2	Q5JJU5_ORYSA	Q5jju5 oryza sativ
34	44	52.4	295	1	ACTB_DICDI	P24005 dictyosteli
35	44	52.4	295	2	Q54XA5_DICDI	Q54xa5 dictyosteli
36	44	52.4	297	2	Q7QB91_ANOGA	Q7qb91 anopheles g
37	44	52.4	317	2	Q4K2D4_STRPN	Q4k2d4 streptococc
38	44	52.4	317	2	Q4K2F0_STRPN	Q4k2f0 streptococc
39	44	52.4	398	2	Q9PNJ8_CAMJE	Q9pnj8 campylobact
40	44	52.4	399	2	Q4HPR6_CAMUP	Q4hpr6 campylobact
41	44	52.4	403	2	Q5HU08_CAMJR	Q5hu08 campylobact
42	44	52.4	446	2	Q4I312_GIBZE	Q4i312 gibberella
43	44	52.4	494	2	Q88TC3_LACPL	Q88tc3 lactobacill
44	44	52.4	515	2	Q57VB3_9TRYP	Q57vb3 trypanosoma
45	44	52.4	571	2	Q54121_STAEP	Q54121 staphylococ

ALIGNMENTS

RESULT 1
CDSN_HUMAN
ID CDSN_HUMAN STANDARD; PRT; 529 AA.
AC Q15517; O43509; Q5SQ85; Q5STD2; Q7LA70; Q7LA71; Q86Z04; Q8IZU4;
AC Q8IZU5; Q8IZU6; Q8N5P3; Q95IF9; Q9NP52; Q9NPE0; Q9NPG5; Q9NRH4;
AC Q9NRH5; Q9NRH6; Q9NRH7; Q9NRH8; Q9UBH8; Q9UIN6; Q9UIN7; Q9UIN8;
AC Q9UIN9; Q9UIP0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Corneodesmosin precursor (S protein).
GN Name=CDSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LEU-18.
RC TISSUE=Foreskin;
RX MEDLINE=94022396; PubMed=8415725;
RA Zhou Y., Chaplin D.D.;
RT "Identification in the HLA class I region of a gene expressed late in
RT keratinocyte differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9470-9474(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT ALA-408.
RC TISSUE=Epidermis;
RX MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT "Expression cloning of human corneodesmosin proves its identity with
RT the product of the S gene and allows improved characterization of its
RT processing during keratinocyte differentiation.";
RL J. Biol. Chem. 273:22640-22647(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS LEU-18; SER-153 DEL;
RP PHE-202 AND LEU-410.
RX MEDLINE=22256218; PubMed=12366786;
RX DOI=10.1034/j.1399-0039.2002.600110.x;
RA Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT patients with psoriasis.";
RL Tissue Antigens 60:77-83(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RT "Genome diversity in HLA: a new strategy for detection of genetic
RT polymorphisms in expressed genes within the HLA class III and class I

```
RT   regions.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANTS LEU-18;
RP   ASN-143; SER-153 DEL; PHE-202; ALA-408; LEU-410 AND ASN-527.
RX   MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA   Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA   Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA   Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA   Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA   Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA   Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA   Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA   Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA   Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA   Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA   Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA   Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA   Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA   Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA   Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA   Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcott R.,
RA   Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA   Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA   Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA   Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA   Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA   Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLay K.,
RA   McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA   Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA   Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA   Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA   Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA   Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA   Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA   Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,
RA   Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA   Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA   Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT   "The DNA sequence and analysis of human chromosome 6.";
RL   Nature 425:805-811(2003).
RN   [7]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS LEU-18; SER-153
RP   DEL; PHE-202 AND LEU-410.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 30-529 (ALLELES 1.11; 1.21; 1.31;
RP   1.32; 1.41; 1.42; 1.43; 1.51; 1.52; 2.11; 2.21; 2.22 AND 2.23), AND
RP   VARIANTS PHE-56; ASN-143 DEL; ASN-150; PHE-202; GLY-401; ALA-408;
RP   LEU-410 AND ASN-527.
RX   MEDLINE=21108612; PubMed=11169256;
RX   DOI=10.1034/j.1399-0039.2001.057001032.x;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [9]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 106-529, VARIANTS ASN-143;
RP   PHE-202; GLY-401; ALA-408 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX   MEDLINE=20066866; PubMed=10599883;
RX   DOI=10.1034/j.1399-0039.1999.540501.x;
RA   Jenisch S., Koch S., Henseler T., Nair R.P., Elder J.T., Watts C.E.,
RA   Westphal E., Voorhees J.J., Christophers E., Kroenke M.;
RT   "Corneodesmosin gene polymorphism demonstrates strong linkage
RT   disequilibrium with HLA and association with psoriasis vulgaris.";
RL   Tissue Antigens 54:439-449(1999).
RN   [10]
RP   PARTIAL PROTEIN SEQUENCE, AND CHARACTERIZATION.
RX   MEDLINE=98058975; PubMed=9395522; DOI=10.1074/jbc.272.50.31770;
RA   Simon M., Montezin M., Guerrin M., Durieux J.-J., Serre G.;
RT   "Characterization and purification of human corneodesmosin, an
RT   epidermal basic glycoprotein associated with corneocyte-specific
RT   modified desmosomes.";
RL   J. Biol. Chem. 272:31770-31776(1997).
RN   [11]
RP   INVOLVEMENT IN HTSS.
RX   PubMed=12754508; DOI=10.1038/ng1163;
RA   Levy-Nissenbaum E., Betz R.C., Frydman M., Simon M., Lahat H.,
RA   Bakhan T., Goldman B., Bygum A., Pierick M., Hillmer A.M., Jonca N.,
RA   Toribio J., Kruse R., Dewald G., Cichon S., Kubisch C., Guerrin M.,
RA   Serre G., Nothen M.M., Pras E.;
RT   "Hypotrichosis simplex of the scalp is associated with nonsense
RT   mutations in CDSN encoding corneodesmosin.";
RL   Nat. Genet. 34:151-153(2003).
RN   [12]
RP   VARIANTS PHE-202 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX   PubMed=10844560; DOI=10.1046/j.1523-1747.2000.00002.x;
RA   Enerbaeck C., Enlund F., Inerot A., Samuelsson L., Wahlstroem J.,
RA   Swanbeck G., Martinsson T.;
RT   "S gene (Corneodesmosin) diversity and its relationship to psoriasis;
RT   high content of cSNP in the HLA-linked S gene.";
RL   J. Invest. Dermatol. 114:1158-1163(2000).
RN   [13]
RP   VARIANTS PHE-202; GLY-401; ALA-408; LEU-410 AND ASN-527, AND
RP   ASSOCIATION WITH PSORIASIS.
RX   PubMed=12472658; DOI=10.1034/j.1399-0039.2002.600403.x;
RA   Orru S., Giuressi E., Casula M., Loizedda A., Murru R., Mulargia M.,
RA   Masala M.V., Cerimele D., Zucca M., Aste N., Biggio P., Carcassi C.,
RA   Contu L.;
RT   "Psoriasis is associated with a SNP haplotype of the corneodesmosin
RT   gene (CDSN).";
RL   Tissue Antigens 60:292-298(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation.
CC   -!- TISSUE SPECIFICITY: Exclusively expressed in skin.
CC   -!- POLYMORPHISM: Genetic variation in CDSN may be associated with
CC       susceptibility to psoriasis [MIM:177900]. Various CDSN alleles are
CC       known including alleles 1.11, 1.21, 1.31, 1.32, 1.41, 1.42, 1.43,
CC       1.51, 1.52, 2.11, 2.21, 2.22 and 2.23.
CC   -!- DISEASE: Defects in CDSN are a cause of hypotrichosis simplex of
CC       the scalp (HTSS) [MIM:146520]; also known as hypotrichosis Spanish
CC       type. HTSS is an autosomal dominant form of isolated alopecia.
CC       Affected individuals have normal hair in early childhood but
CC       experience progressive loss of scalp hair beginning in the middle
CC       of the first decade and almost complete baldness by the third
CC       decade.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 501.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; L20815; AAA21321.1; ALT_SEQ; mRNA.
DR   EMBL; AF030130; AAC24196.1; -; mRNA.

  Query Match             92.9%;  Score 78;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 2
CDSN_PANTR
ID   CDSN_PANTR     STANDARD;      PRT;   529 AA.
AC   Q7YR44;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BA000041; BAC78170.1; -; Genomic_DNA.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    529       Corneodesmosin.
FT   COMPBIAS     62    464       Ser-rich.
FT   CARBOHYD    172    172       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;

  Query Match             92.9%;  Score 78;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 3
CDSN_MACMU
ID   CDSN_MACMU     STANDARD;      PRT;   534 AA.
AC   Q5TM45;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=CDSN;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15269276; DOI=10.1093/molbev/msh216;
RA   Kulski J.K., Anzai T., Shiina T., Inoko H.;
RT   "Rhesus macaque class I duplicon structures, organization, and
RT   evolution within the alpha block of the major histocompatibility
RT   complex.";
RL   Mol. Biol. Evol. 21:2079-2091(2004).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB128049; BAD69749.1; -; Genomic_DNA.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    534       Corneodesmosin.
FT   COMPBIAS     62    469       Ser-rich.
SQ   SEQUENCE   534 AA;  51983 MW;  CB6A9DF256C32874 CRC64;

  Query Match             92.9%;  Score 78;  DB 1;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.0015;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        296 KSYGGYEVVGGSSDS 310


RESULT 4
CDSN_MOUSE
ID   CDSN_MOUSE     STANDARD;      PRT;   561 AA.
AC   Q7TPC1;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=Cdsn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C3H/He; TISSUE=Osteoblast;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

```
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BC055373; AAH55373.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039518; Mus musculus.
DR   MGI; MGI:3505689; Cdsn.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    561       Corneodesmosin.
FT   COMPBIAS     62    493       Ser-rich.
SQ   SEQUENCE   561 AA;  54414 MW;  7A6EBEF364038746 CRC64;

  Query Match             77.4%;  Score 65;  DB 1;  Length 561;
  Best Local Similarity   80.0%;  Pred. No. 0.17;
  Matches   12;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              | |||||||||::|
Db        318 KPYGGYEVVGGSANS 332


RESULT 5
CDSN_PIG
ID   CDSN_PIG       STANDARD;      PRT;   225 AA.
AC   O19084;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin (S protein) (Fragment).
GN   Name=CDSN;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   TISSUE=Peripheral blood;
RA   Velten F.W., Rogel-Gaillard C., Renard C., Pontarotti P.,
RA   Tazi-Ahnini R., Vaiman M., Chardon P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; Z97402; CAB10690.1; -; Genomic_DNA.
FT   COMPBIAS     79     85       Poly-Ser.
FT   COMPBIAS    129    140       Poly-Ser.
FT   COMPBIAS    159    168       Poly-Ser.
FT   COMPBIAS    179    184       Poly-Ser.
FT   NON_TER       1      1
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  21644 MW;  8544070C41F7B280 CRC64;

  Query Match             70.2%;  Score 59;  DB 1;  Length 225;
  Best Local Similarity   84.6%;  Pred. No. 0.54;
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSDS 15
              :| |||||||||
Db         13 FGSYEVVGGSSDS 25


RESULT 6
Q6NUU9_BRARE
ID   Q6NUU9_BRARE PRELIMINARY;      PRT;   293 AA.
AC   Q6NUU9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Phosphatidylinositol glycan, class C.
GN   Name=pigc; ORFNames=DKEY-19F21.4-001;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RG   NIH MGC Project;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Johnson C.;
RL   Submitted (DEC-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068421; AAH68421.1; -; mRNA.
DR   EMBL; BX470128; CAI20802.1; -; Genomic_DNA.
DR   ZFIN; ZDB-GENE-030131-2714; pigc.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0017176; F:phosphatidylinositol N-acetylglucosaminyltr. . .; IEA.
DR   GO; GO:0006506; P:GPI anchor biosynthesis; IEA.
DR   InterPro; IPR009450; GPI2.
DR   Pfam; PF06432; GPI2; 1.
SQ   SEQUENCE   293 AA;  32420 MW;  3613FE272E8F40C6 CRC64;

  Query Match             58.3%;  Score 49;  DB 2;  Length 293;
  Best Local Similarity   75.0%;  Pred. No. 27;
  Matches    9;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              || |||:||:||
Db         98 YGLYEVLGGASD 109
```

```
RESULT 7
AROE_METTH
ID   AROE_METTH     STANDARD;      PRT;   283 AA.
AC   O26344;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Shikimate dehydrogenase (EC 1.1.1.25).
GN   Name=aroE; OrderedLocusNames=MTH242;
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 3-dehydroshikimate +
CC       NADPH.
CC   -!- PATHWAY: Metabolic intermediate biosynthesis; chorismate
CC       biosynthesis; chorismate from D-erythrose 4-phosphate and PEP:
CC       step 4.
CC   -!- PATHWAY: Context: Aromatic amino acids biosynthesis.
CC   -!- SIMILARITY: Belongs to the shikimate dehydrogenase family.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 17.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AE000811; AAB84748.1; ALT_FRAME; Genomic_DNA.
DR   HSSP; Q58484; 1NVT.
DR   HAMAP; MF_00222; -; 1.
DR   InterPro; IPR011342; AroE.
DR   InterPro; IPR006151; Shikimate_DH.
DR   Pfam; PF01488; Shikimate_DH; 1.
DR   TIGRFAMs; TIGR00507; aroE; 1.
KW   Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
KW   Complete proteome; NADP; Oxidoreductase.
SQ   SEQUENCE   283 AA;  30484 MW;  3293E04F92FC0B89 CRC64;

  Query Match             57.1%;  Score 48;  DB 1;  Length 283;
  Best Local Similarity   64.3%;  Pred. No. 37;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDS 15
              ||||||:: ||  |
Db        175 SYGGYELIQGSVKS 188


RESULT 8
Q6YXZ0_ORYSA
ID   Q6YXZ0_ORYSA PRELIMINARY;      PRT;   367 AA.
AC   Q6YXZ0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein OSJNBa0026E05.27.
GN   Name=OSJNBa0026E05.27;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005647; BAD10537.1; -; Genomic_DNA.
DR   Gramene; Q6YXZ0; -.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
KW   Hypothetical protein.
SQ   SEQUENCE   367 AA;  41340 MW;  1A6118B4831906DD CRC64;

  Query Match             57.1%;  Score 48;  DB 2;  Length 367;
  Best Local Similarity   75.0%;  Pred. No. 49;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDS 15
              ||| ||||  ||
Db        191 GGYAVVGGGGDS 202


RESULT 9
Q4HZJ7_GIBZE
ID   Q4HZJ7_GIBZE PRELIMINARY;      PRT;  1473 AA.
AC   Q4HZJ7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG09611.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000397; EAA76233.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1473 AA;  164340 MW;  6576C9D91C24A502 CRC64;

  Query Match             56.0%;  Score 47;  DB 2;  Length 1473;
  Best Local Similarity   69.2%;  Pred. No. 3.2e+02;
```

```
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDSG 16
              ||  |:|||| ||
Db        610 GGLNVIGGSSGSG 622


RESULT 10
Q9DDJ7_CHICK
ID   Q9DDJ7_CHICK PRELIMINARY;      PRT;   115 AA.
AC   Q9DDJ7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Retinoblastoma tumor suppressor (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21134552; PubMed=11241829;
RX   DOI=10.1002/1097-0177(20010301)220:3<198::AID-DVDY1103>3.0.CO;2-T;
RA   Wagner M., Miles K., Siddiqui M.A.Q.;
RT   "Early developmental expression pattern of retinoblastoma tumor
RT   suppressor mRNA indicates a role in the epithelial-to-mesenchyme
RT   transformation of endocardial cushion cells.";
RL   Dev. Dyn. 220:198-211(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Wagner M.A.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF323706; AAG45219.1; -; mRNA.
DR   Ensembl; ENSGALG00000011597; Gallus gallus.
FT   NON_TER       1      1
SQ   SEQUENCE   115 AA;  12680 MW;  CE688F50FD759136 CRC64;

  Query Match             54.8%;  Score 46;  DB 2;  Length 115;
  Best Local Similarity   72.7%;  Pred. No. 28;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              ::|||||| ||
Db         12 EAYGGYEVDGG 22


RESULT 11
TXPH_PYETR
ID   TXPH_PYETR     STANDARD;      PRT;   289 AA.
AC   Q07338;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Toxin tox21A precursor (Insect-selective neurotoxin TxP-I homolog).
OS   Pyemotes tritici (Straw itch mite).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Acariformes; Trombidiformes; Prostigmata; Anystina; Eleutherengona;
OC   Heterostigmata; Pyemotoidea; Pyemotidae; Pyemotes.
OX   NCBI_TaxID=6950;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93227269; PubMed=8470135; DOI=10.1016/0041-0101(93)90149-D;
RA   Tomalski M.D., Hutchinson K., Todd J., Miller L.K.;
RT   "Identification and characterization of tox21A: a mite cDNA encoding a
RT   paralytic neurotoxin related to TxP-I.";
RL   Toxicon 31:319-326(1993).
CC   -!- FUNCTION: Has contracting-paralyzing activity in insect larvae.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Posterior glands which appear to be connected
CC       with the stylet through a series of ducts (Probable).
CC   -!- PTM: Contains several disulfide bonds (By similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; S58065; AAB26160.1; -; mRNA.
DR   EMBL; L09272; AAA29801.1; -; mRNA.
DR   PIR; A45369; A45369.
KW   Multigene family; Neurotoxin; Signal; Toxin.
FT   SIGNAL        1     14       Potential.
FT   PROPEP       15     27       By similarity.
FT   CHAIN        28    289       Toxin tox21A.
SQ   SEQUENCE   289 AA;  32732 MW;  476534ED1B84CB5B CRC64;

  Query Match             54.8%;  Score 46;  DB 1;  Length 289;
  Best Local Similarity   46.2%;  Pred. No. 78;
  Matches    6;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSS 13
              | |||:::::|| :
Db         80 KVYGGFDIIGGKA 92


RESULT 12
Q857A1_9CAUD
ID   Q857A1_9CAUD PRELIMINARY;      PRT;   387 AA.
AC   Q857A1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Gp33.
OS   Mycobacteriophage Rosebush.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX   NCBI_TaxID=205874;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA   Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA   Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA   Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA   Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA   Hatfull G.F.;
RT   "Origins of highly mosaic mycobacteriophage genomes.";
RL   Cell 113:171-182(2003).
DR   EMBL; AY129334; AAN01875.1; -; Genomic_DNA.
SQ   SEQUENCE   387 AA;  38655 MW;  0F4669244C14EE58 CRC64;

  Query Match             54.8%;  Score 46;  DB 2;  Length 387;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+02;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDSG 16
              || |   ||| | ||
Db        332 SYNGISAVGGKSSSG 346


RESULT 13
Q8XLS8_CLOPE
ID   Q8XLS8_CLOPE PRELIMINARY;      PRT;   540 AA.
AC   Q8XLS8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Probable transporter.
GN   OrderedLocusNames=CPE0963;
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=13 / Type A;
```

```
RX   MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; BA000016; BAB80669.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR   GO; GO:0006865; P:amino acid transport; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR002293; AA/rel_permease1.
DR   InterPro; IPR004841; Permease_region.
DR   Pfam; PF00324; AA_permease; 1.
KW   Complete proteome; Transmembrane; Transport.
SQ   SEQUENCE   540 AA;  60425 MW;  D89A82D3B91661ED CRC64;

  Query Match             54.8%;  Score 46;  DB 2;  Length 540;
  Best Local Similarity   69.2%;  Pred. No. 1.5e+02;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSD 14
              :||| |||||  |
Db        223 AYGGIEVVGGLVD 235


RESULT 14
Q7VF32_HELHP
ID   Q7VF32_HELHP PRELIMINARY;      PRT;   682 AA.
AC   Q7VF32;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Iron-regulated outer membrane protein FrpB.
GN   Name=frpB; OrderedLocusNames=HH1847;
OS   Helicobacter hepaticus.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=32025;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 51449 / 3B1;
RX   MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA   Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA   Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA   Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA   Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT   "The complete genome sequence of the carcinogenic bacterium
RT   Helicobacter hepaticus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR   EMBL; AE017149; AAP78444.1; -; Genomic_DNA.
DR   GO; GO:0019867; C:outer membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000531; TonB_receptor.
DR   Pfam; PF00593; TonB_dep_Rec; 1.
KW   Complete proteome; Membrane; Outer membrane; TonB box.
SQ   SEQUENCE   682 AA;  76159 MW;  ECF2364CCF1BFA05 CRC64;

  Query Match             54.8%;  Score 46;  DB 2;  Length 682;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    8;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              :||| : |:|| |: |
Db        163 RSYGAHFVLGGQSNKG 178


RESULT 15
Q9I3U9_PSEAE
ID   Q9I3U9_PSEAE PRELIMINARY;      PRT;   138 AA.
AC   Q9I3U9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PA1395;
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004569; AAG04784.1; -; Genomic_DNA.
DR   PIR; E83470; E83470.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   138 AA;  15391 MW;  3CC352262D709A19 CRC64;

  Query Match             53.6%;  Score 45;  DB 2;  Length 138;
  Best Local Similarity   69.2%;  Pred. No. 50;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSD 14
              ||  ||||| ||:
Db         69 SYSRYEVVGSSSE 81


Search completed: March 21, 2006, 21:48:44
Job time : 50.4182 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:43 ; Search time 51.4909 Seconds
(without alignments)
136.530 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	92.9	486	7	ABJ39147	Abj39147 Human S p
2	78	92.9	486	8	ADQ21504	Adq21504 Human sof
3	78	92.9	529	2	AAW80498	Aaw80498 A protein
4	78	92.9	529	2	AAW85055	Aaw85055 Human pro
5	78	92.9	529	3	AAY44989	Aay44989 Human epi
6	78	92.9	529	4	AAG64152	Aag64152 Polypepti
7	78	92.9	529	4	AAU27550	Aau27550 Corneodes
8	78	92.9	529	7	ABJ39148	Abj39148 Human cor
9	78	92.9	529	8	ADN04226	Adn04226 Antipsori
10	78	92.9	529	8	ADP12525	Adp12525 Protein e
11	48	57.1	260	8	ADN18469	Adn18469 Bacterial
12	46	54.8	81	8	ADT58307	Adt58307 Plant pol
13	46	54.8	289	2	AAR23623	Aar23623 Tox2a gen
14	46	54.8	289	2	AAW56745	Aaw56745 Insect-sp
15	46	54.8	289	4	AAE02823	Aae02823 Pyemotes
16	45	53.6	143	7	ABO75700	Abo75700 Pseudomon
17	45	53.6	173	4	AAB76614	Aab76614 Corynebac
18	45	53.6	291	2	AAR23622	Aar23622 Tox34 gen
19	45	53.6	291	2	AAW19987	Aaw19987 Pyemotes
20	45	53.6	291	2	AAW56744	Aaw56744 Insect-sp
21	45	53.6	291	4	AAE02822	Aae02822 Pyemotes
22	45	53.6	438	4	AAB76612	Aab76612 Corynebac
23	45	53.6	513	6	ABU24396	Abu24396 Protein e
24	45	53.6	537	2	AAW69353	Aaw69353 B. lactof
25	45	53.6	537	4	AAG90530	Aag90530 C glutami
26	45	53.6	537	4	AAB83171	Aab83171 Corynebac
27	45	53.6	894	7	ABO72380	Abo72380 Pseudomon
28	44	52.4	385	3	AAY95075	Aay95075 Candida a
29	44	52.4	398	6	ABU26583	Abu26583 Protein e
30	44	52.4	581	8	ADS28153	Ads28153 Bacterial
31	43	51.2	113	8	ADS30231	Ads30231 Bacterial
32	43	51.2	208	4	AAU34330	Aau34330 Staphyloc
33	43	51.2	208	6	ABU16428	Abu16428 Protein e
34	43	51.2	210	4	AAU37379	Aau37379 Staphyloc
35	43	51.2	210	4	AAU37481	Aau37481 Staphyloc
36	43	51.2	213	6	ABM72100	Abm72100 Staphyloc
37	43	51.2	334	8	ADS30335	Ads30335 Bacterial
38	43	51.2	361	9	ABM90889	Abm90889 M. xanthu
39	43	51.2	508	4	ABG09910	Abg09910 Novel hum
40	43	51.2	539	8	ADS28025	Ads28025 Bacterial
41	43	51.2	782	7	ADD25528	Add25528 Binding d
42	43	51.2	1266	8	ADY24617	Ady24617 Plant ful
43	43	51.2	1625	8	ADU07647	Adu07647 Amino aci
44	42	50.0	53	6	ABR98377	Abr98377 Tumour ce
45	42	50.0	53	7	ADC84723	Adc84723 MCF-7 bre

ALIGNMENTS

RESULT 1
ABJ39147
ID ABJ39147 standard; protein; 486 AA.
XX
AC ABJ39147;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human S protein encoded by the skin corneodesmosin (CDSN) gene (S-gene).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW S protein; skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43848.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to

```
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human S protein (corneodesmosin protein only
CC   expressed in the skin), the target of the invention
XX
SQ   Sequence 486 AA;

  Query Match             92.9%;  Score 78;  DB 7;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        275 KSYGGYEVVGGSSDS 289


RESULT 2
ADQ21504
ID   ADQ21504 standard; protein; 486 AA.
XX
AC   ADQ21504;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 4324.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 4324; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 486 AA;

  Query Match             92.9%;  Score 78;  DB 8;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        275 KSYGGYEVVGGSSDS 289


RESULT 3
AAW80498
ID   AAW80498 standard; protein; 529 AA.
XX
AC   AAW80498;
XX
DT   10-FEB-1999  (first entry)
XX
DE   A protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761362-A1.
XX
PD   02-OCT-1998.
XX
PF   28-MAR-1997;   97FR-00003899.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586093/50.
DR   N-PSDB; AAV63614.
XX
PT   New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT   or therapeutic treatment of trophic disorders or thinning of the
PT   epidermis.
XX
PS   Claim 1; Page 29; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

  Query Match             92.9%;  Score 78;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 4
AAW85055
ID   AAW85055 standard; protein; 529 AA.
XX
AC   AAW85055;
XX
DT   05-FEB-1999  (first entry)
```

XX
DE Human protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761363-A1.
XX
PD 02-OCT-1998.
XX
PF 11-SEP-1997; 97FR-00011317.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586094/50.
DR N-PSDB; AAV63541.
XX
PT New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT epidermis etc., also compositions containing protease that degrades
PT corneodesmosin.
XX
PS Claim 1; Page 30; 35pp; French.
XX
CC The present sequence represents a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products
XX
SQ Sequence 529 AA;

Query Match 92.9%; Score 78; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSYGGYEVVGGSSDS 15
||||||||||||||||
Db 291 KSYGGYEVVGGSSDS 305

RESULT 5
AAY44989
ID AAY44989 standard; protein; 529 AA.
XX
AC AAY44989;
XX
DT 23-MAY-2000 (first entry)
XX
DE Human epidermal protein-6.
XX
KW Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW dermatological; anitarteriosclerotic; antiinflammatory;
KW immunosuppressive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .32
FT /label= Signal_peptide
FT Modified-site 3
FT /note= "Potential phosphorylation site"
FT Modified-site 37
FT /note= "Potential phosphorylation site"
FT Modified-site 50
FT /note= "Potential phosphorylation site"
FT Modified-site 57
FT /note= "Potential phosphorylation site"
FT Modified-site 80
FT /note= "Potential phosphorylation site"
FT Modified-site 105
FT /note= "Potential phosphorylation site"
FT Modified-site 172
FT /note= "N-glycosylated"
FT Modified-site 182
FT /note= "Potential phosphorylation site"
FT Modified-site 287
FT /note= "Potential phosphorylation site"
FT Modified-site 395
FT /note= "Potential phosphorylation site"
FT Modified-site 440
FT /note= "Potential phosphorylation site"
FT Modified-site 498
FT /note= "Potential phosphorylation site"
FT Modified-site 498
FT /note= "Potential phosphorylation site"
XX
PN WO200006727-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US017107.
XX
PR 28-JUL-1998; 98US-0155203P.
PR 07-DEC-1998; 98US-0155254P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Lal P, Corley NC, Guegler KJ, Patterson C, Baughn MR;
PI Yue H;
XX
DR WPI; 2000-195295/17.
DR N-PSDB; AAZ50582.
XX
PT New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT diagnosis, treatment and prevention of epithelial, cell proliferative,
PT and autoimmune inflammatory disorders.
XX
PS Claim 1; Fig 4; 82pp; English.
XX
CC The present sequence is human epidermal protein-6 (HEPI) expressed in
CC breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC cDNA library. HEPI-6 shares 99% identity with human S protein.
CC Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC for protein expression. The HEPI proteins are useful for the treatment of
CC epithelial disorders, including dyshidrotic eczema and scabies, cell
CC proliferative disorders including actinic keratosis and arteriosclerosis,
CC and autoimmune/inflammatory disorders like acquired immune deficiency
CC syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC comprising HEPI proteins are useful for treating disorders associated
CC with altered HEPI expression
XX
SQ Sequence 529 AA;

Query Match 92.9%; Score 78; DB 3; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSYGGYEVVGGSSDS 15
||||||||||||||||

```
Db        291 KSYGGYEVVGGSSDS 305


RESULT 6
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 143
FT                   /label= Ser, Asn
FT   Misc-difference 151
FT                   /label= Ser, Cys
FT   Misc-difference 152
FT                   /label= Ser, Arg
FT   Misc-difference 200
FT                   /label= Ser, Phe
FT   Misc-difference 282
FT                   /label= Pro, Thr
FT   Misc-difference 401
FT                   /label= Ser, Gly
FT   Misc-difference 408
FT                   /label= Ala, Ser
FT   Misc-difference 409
FT                   /label= Gly, Val
FT   Misc-difference 410
FT                   /label= Ser, Leu
FT   Misc-difference 453
FT                   /label= Ser, Asn
FT   Misc-difference 527
FT                   /label= Asp, Asn
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

  Query Match             92.9%;  Score 78;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 7
AAU27550
ID   AAU27550 standard; protein; 529 AA.
XX
AC   AAU27550;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
DR   N-PSDB; AAS43749.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Claim 8; Fig 2; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. The present sequence represents corneodesmosin amino acid
CC   sequence, as described in the invention
XX
SQ   Sequence 529 AA;

  Query Match             92.9%;  Score 78;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305
```

```
RESULT 8
ABJ39148
ID   ABJ39148 standard; protein; 529 AA.
XX
AC   ABJ39148;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human corneodesmosin (CDSN) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 28
FT                   /note= "Encoded by CCAG"
FT   misc_difference 408
FT                   /note= "Encoded by TCC"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   N-PSDB; ABT43849, ABT43850.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
XX
SQ   Sequence 529 AA;

  Query Match             92.9%;  Score 78;  DB 7;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305
```

```
RESULT 9
ADN04226
ID   ADN04226 standard; protein; 529 AA.
XX
AC   ADN04226;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #308.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04225.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 529 AA;

  Query Match             92.9%;  Score 78;  DB 8;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0013;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305
```

```
RESULT 10
ADP12525
ID   ADP12525 standard; protein; 529 AA.
XX
AC   ADP12525;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #135.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 529 AA;

Query Match 92.9%; Score 78; DB 8; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSYGGYEVVGGSSDS 15
||||||||||||||||
Db 291 KSYGGYEVVGGSSDS 305

RESULT 11
ADN18469
ID ADN18469 standard; protein; 260 AA.
XX
AC ADN18469;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #1122.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 1122; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 260 AA;

Query Match 57.1%; Score 48; DB 8; Length 260;
Best Local Similarity 64.3%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SYGGYEVVGGSSDS 15
|||||:: || |
Db 152 SYGGYELIQGSVKS 165

RESULT 12
ADT58307
ID ADT58307 standard; protein; 81 AA.
XX
AC ADT58307;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant polypeptide, SEQ ID 8384.
XX
KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.

XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
DR WPI; 2004-757369/74.
XX
PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
PS Claim 2; SEQ ID NO 8384; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 81 AA;

Query Match 54.8%; Score 46; DB 8; Length 81;
Best Local Similarity 43.8%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSSDSG 16
              | |||:|   |: : |
Db         50 KGYGGFEAAAGAENGG 65
```

RESULT 13
AAR23623
ID AAR23623 standard; protein; 289 AA.
XX
AC AAR23623;
XX
DT 25-MAR-2003 (revised)
DT 30-OCT-1992 (first entry)
XX
DE Tox2a gene product.
XX
KW P. tritici; Baculovirus; Galleria mellonella; insect control agents;
KW arthropods.
XX
OS Pyemotes tritici.
XX
PN WO9206181-A.
XX
PD 16-APR-1992.
XX
PF 01-OCT-1991; 91WO-US007216.
XX
PR 04-OCT-1990; 90US-00593657.
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
PI Tomalski MD, Miller LK;
XX
DR WPI; 1992-150868/18.
DR N-PSDB; AAQ24177.
XX
PT DNA encoding insect-specific paralytic neurotoxin(s) - for use in
PT biological insect control agents with improved efficiency.
XX
PS Disclosure; Page 77; 101pp; English.
XX
CC This sequence is an insect specific paralytic neurotoxic gene from P.
CC tritici designated Tox2a. Nucleotides 118-873 from the corresponding DNA
CC may be use within a recombinant DNA molecule which encodes an insect
CC paralytic neurotoxin. P. tritici toxins are not acutely toxic to mammals
CC but when tested on Galleria mellonella (wax moth) larvae cause rapid
CC muscle-contracting paralysis. This sequence can be transformed in a
CC Baculovirus host, as Baculovirus' only infect arthropods. These
CC neurotoxins have a primary use as genetically engineered insect control
CC agents within agricultural compsn. for applying to plants, plant
CC environments or distributed in baits to effect the biological control of
CC insect pests. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 289 AA;

Query Match 54.8%; Score 46; DB 2; Length 289;
Best Local Similarity 46.2%; Pred. No. 72;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| : :
Db         80 KVYGGFDIIGGKA 92
```

RESULT 14
AAW56745
ID AAW56745 standard; protein; 289 AA.
XX
AC AAW56745;
XX
DT 02-SEP-1998 (first entry)
XX
DE Insect-specific paralytic neurotoxin Tox21a from Pyemotes tritici.
XX
KW Insect-specific paralytic neurotoxin; Tox21a; straw-itch mite;
KW ventricosus group; baculovirus promoter; heat shock gene promoter;
KW baculovirus; Autographa californica nuclear polyhedrosis virus;
KW Helicoverpa zea nucleopolyhedrosis virus;
KW ecdysteroid UDP-glycosyl transferase; insect control.
XX
OS Pyemotes tritici.
XX
PN WO9814578-A1.
XX
PD 09-APR-1998.
XX

PF 01-OCT-1997; 97WO-US018077.
XX
PR 01-OCT-1996; 96US-00720606.
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Miller L, Lu A, Black BC, Dierks M;
XX
DR WPI; 1998-240084/21.
DR N-PSDB; AAV29479.
XX
PT New recombinant DNA for control of insects - comprising insect-specific
PT toxin DNA under the control of a baculovirus 6.9 or DA26 promoter or a
PT heat shock gene promoter.
XX
PS Disclosure; Page 53; 83pp; English.
XX
CC The present sequence represents an insect-specific paralytic neurotoxin
CC called Tox21a. The toxic protein is produced by Pyemotes tritici, the
CC straw-itch mite of the ventricosus group of the Pyemotes. The toxin is
CC used to produce the construct of the invention. This comprises a promoter
CC selected from a baculovirus 6.9K promoter, a baculovirus DA26 promoter
CC and a heat shock gene promoter, and a coding sequence for an insect-
CC specific toxin, the coding sequence being operably linked to and
CC expressed under the regulatory control of the promoter. Baculoviruses,
CC particularly Autographa californica nuclear polyhedrosis virus or
CC Helicoverpa zea nucleopolyhedrosis virus where the gene encoding
CC ecdysteroid UDP-glycosyl transferase has been inactivated, can also be
CC engineered to contain the toxin sequence. The compositions are used for
CC the control of insect pests on plants
XX
SQ Sequence 289 AA;

Query Match 54.8%; Score 46; DB 2; Length 289;
Best Local Similarity 46.2%; Pred. No. 72;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| :
Db         80 KVYGGFDIIGGKA 92
```

RESULT 15
AAE02823
ID AAE02823 standard; protein; 289 AA.
XX
AC AAE02823;
XX
DT 06-AUG-2001 (first entry)
XX
DE Pyemotes tritici paralytic neurotoxin Tox21a.
XX
KW Recombinant baculovirus; insect-specific toxin; paralytic neurotoxin;
KW biological pest control; insect pest; Tox21a; mite.
XX
OS Pyemotes tritici.
XX
PN US6235278-B1.
XX
PD 22-MAY-2001.
XX
PF 01-OCT-1997; 97US-00942012.
XX
PR 01-OCT-1996; 96US-00720606.
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Miller LK, Lu A, Black BC, Dierks PM;
XX
DR WPI; 2001-366397/38.
DR N-PSDB; AAD07227.
XX
PT New recombinant baculoviruses expressing insect-specific toxins,
PT specifically paralytic neurotoxin from Pyemotes, useful as biological
PT control agents in controlling insect pests.
XX
PS Claim 4; Col 41-44; 58pp; English.
XX
CC The invention relates to a recombinant baculovirus containing a
CC baculovirus 6.9K promoter, a baculovirus DA26 promoter or a heat shock
CC gene promoter, and a coding sequence for an insect-specific toxin,
CC preferably a paralytic neurotoxin of a mite of the genus Pyemotes. The
CC recombinant baculovirus is useful in insect toxin compositions, or as
CC biological insect control agent. Compositions comprising the baculovirus
CC can be used in the control of insect pests and in protecting plants from
CC insect pests by reducing feeding and killing susceptible insects. The
CC present sequence is Pyemotes tritici Tox21a protein which is an insect-
CC specific paralytic neurotoxin
XX
SQ Sequence 289 AA;

Query Match 54.8%; Score 46; DB 4; Length 289;
Best Local Similarity 46.2%; Pred. No. 72;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| :
Db         80 KVYGGFDIIGGKA 92
```

Search completed: March 21, 2006, 21:45:55
Job time : 54.4909 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:43 ; Search time 41.8909 Seconds
(without alignments)
159.588 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	92.9	486	4	US-10-316-175-12	Sequence 12, Appl
2	78	92.9	486	5	US-10-723-860-4324	Sequence 4324, Ap
3	78	92.9	486	5	US-10-883-805-16	Sequence 16, Appl
4	78	92.9	529	4	US-10-164-230-61	Sequence 61, Appl
5	78	92.9	529	4	US-10-204-884-3	Sequence 3, Appli
6	78	92.9	529	4	US-10-316-175-15	Sequence 15, Appl
7	78	92.9	529	5	US-10-883-805-6	Sequence 6, Appli
8	78	92.9	529	5	US-10-677-865-1	Sequence 1, Appli
9	48	57.1	107	4	US-10-425-115-354464	Sequence 354464,
10	48	57.1	260	4	US-10-369-493-1122	Sequence 1122, Ap
11	48	57.1	375	4	US-10-437-963-138974	Sequence 138974,
12	46	54.8	81	5	US-10-739-930-8384	Sequence 8384, Ap
13	45	53.6	173	4	US-10-627-476-210	Sequence 210, App
14	45	53.6	438	4	US-10-627-476-206	Sequence 206, App
15	45	53.6	513	4	US-10-282-122A-52320	Sequence 52320, A
16	45	53.6	537	3	US-09-738-626-4284	Sequence 4284, Ap
17	45	53.6	537	6	US-11-073-550-8	Sequence 8, Appli
18	45	53.6	537	6	US-11-073-560-8	Sequence 8, Appli
19	44	52.4	398	4	US-10-282-122A-54507	Sequence 54507, A
20	44	52.4	436	4	US-10-424-599-212211	Sequence 212211,
21	44	52.4	581	4	US-10-369-493-17186	Sequence 17186, A
22	44	52.4	1238	4	US-10-437-963-192399	Sequence 192399,
23	43	51.2	113	4	US-10-369-493-19264	Sequence 19264, A
24	43	51.2	208	3	US-09-815-242-5826	Sequence 5826, Ap
25	43	51.2	208	4	US-10-282-122A-44352	Sequence 44352, A
26	43	51.2	210	3	US-09-815-242-12972	Sequence 12972, A
27	43	51.2	210	3	US-09-815-242-13074	Sequence 13074, A
28	43	51.2	334	4	US-10-369-493-19368	Sequence 19368, A
29	43	51.2	454	4	US-10-767-701-45105	Sequence 45105, A
30	43	51.2	478	4	US-10-437-963-107407	Sequence 107407,
31	43	51.2	508	5	US-10-450-763-40269	Sequence 40269, A
32	43	51.2	539	4	US-10-369-493-17058	Sequence 17058, A
33	43	51.2	782	4	US-10-207-655-89	Sequence 89, Appl
34	43	51.2	955	4	US-10-156-761-15004	Sequence 15004, A
35	43	51.2	1266	4	US-10-425-114-72401	Sequence 72401, A
36	43	51.2	1266	4	US-10-425-115-240496	Sequence 240496,
37	42	50.0	40	4	US-10-437-963-153481	Sequence 153481,
38	42	50.0	48	4	US-10-425-115-314170	Sequence 314170,
39	42	50.0	53	4	US-10-199-820-255	Sequence 255, App
40	42	50.0	59	4	US-10-425-115-290188	Sequence 290188,
41	42	50.0	166	4	US-10-424-599-223245	Sequence 223245,
42	42	50.0	192	4	US-10-437-963-143943	Sequence 143943,
43	42	50.0	193	4	US-10-476-204-13	Sequence 13, Appl
44	42	50.0	199	4	US-10-424-599-199753	Sequence 199753,
45	42	50.0	243	4	US-10-476-204-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-316-175-12
; Sequence 12, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-12

Query Match 92.9%; Score 78; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        275 KSYGGYEVVGGSSDS 289
```

RESULT 2
US-10-723-860-4324
; Sequence 4324, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4324
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-4324

  Query Match             92.9%;  Score 78;  DB 5;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.0029;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        275 KSYGGYEVVGGSSDS 289


RESULT 3
US-10-883-805-16
; Sequence 16, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 16
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No: g414810
US-10-883-805-16

  Query Match             92.9%;  Score 78;  DB 5;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.0029;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        275 KSYGGYEVVGGSSDS 289


RESULT 4
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             92.9%;  Score 78;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 5
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
```

```
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (143)..(143)
;   OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (401)..(401)
;   OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (409)..(409)
;   OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (410)..(410)
;   OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (527)..(527)
;   OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

  Query Match             92.9%;  Score 78;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 6
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-15

  Query Match             92.9%;  Score 78;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 7
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

  Query Match             92.9%;  Score 78;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 8
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
;  APPLICANT: SERRE, Guy Bruno Rene
;  APPLICANT:  SIMON, Michel
```

```
;  APPLICANT:  WEBER-VIVAT, Marina
;  TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
;  TITLE OF INVENTION:  AND USE THEREOF
;  FILE REFERENCE: 016800-336
;  CURRENT APPLICATION NUMBER: US/10/677,865
;  CURRENT FILING DATE:  2003-10-03
;  PRIOR APPLICATION NUMBER: US/09/381,656
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: FR 97/03899
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-10-677-865-1

  Query Match             92.9%;  Score 78;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.0032;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 9
US-10-425-115-354464
; Sequence 354464, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 354464
;   LENGTH: 107
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_86442C.1.pep
US-10-425-115-354464

  Query Match             57.1%;  Score 48;  DB 4;  Length 107;
  Best Local Similarity   56.2%;  Pred. No. 18;
  Matches    9;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              :|:| :|  ||||| |
Db          5 QSFGDFERFGGSSDRG 20


RESULT 10
US-10-369-493-1122
; Sequence 1122, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1122
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Methanobacterium thermoautotrophicum
US-10-369-493-1122

  Query Match             57.1%;  Score 48;  DB 4;  Length 260;
  Best Local Similarity   64.3%;  Pred. No. 45;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDS 15
              ||||||:: ||  |
Db        152 SYGGYELIQGSVKS 165


RESULT 11
US-10-437-963-138974
; Sequence 138974, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 138974
;   LENGTH: 375
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_40311C.1.pep
US-10-437-963-138974

  Query Match             57.1%;  Score 48;  DB 4;  Length 375;
  Best Local Similarity   75.0%;  Pred. No. 66;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDS 15
              ||| ||||  ||
Db        199 GGYAVVGGGGDS 210


RESULT 12
US-10-739-930-8384
; Sequence 8384, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 8384
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Zea mays
```

```
;   FEATURE:
;   OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C5923_19.p
US-10-739-930-8384

  Query Match             54.8%;  Score 46;  DB 5;  Length 81;
  Best Local Similarity   43.8%;  Pred. No. 27;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              | |||:|   |: : |
Db         50 KGYGGFEAAAGAENGG 65


RESULT 13
US-10-627-476-210
; Sequence 210, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Mark
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schoder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
;  TITLE OF INVENTION:  TRANSPORT
;  FILE REFERENCE: BGI-125CPCN
;  CURRENT APPLICATION NUMBER: US/10/627,476
;  CURRENT FILING DATE:  2003-07-25
;  PRIOR APPLICATION NUMBER: 09/602,787
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: USSN 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: DE 19931454.3
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931478.0
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931563.9
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932122.1
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932124.8
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932125.6
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932128.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932180.9
;  PRIOR FILING DATE: 1999-07-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 210
;   LENGTH: 173
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 70
;   OTHER INFORMATION: Xaa indicates any amino acid
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 72
;   OTHER INFORMATION: Xaa indicates any amino acid
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 86
;   OTHER INFORMATION: Xaa indicates any amino acid
US-10-627-476-210

  Query Match             53.6%;  Score 45;  DB 4;  Length 173;
  Best Local Similarity   62.5%;  Pred. No. 83;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              |||||   | || | |
Db        154 KSYGGAYCVMGSKDMG 169


RESULT 14
US-10-627-476-206
; Sequence 206, Application US/10627476
; Publication No. US20040030116A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Mark
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schoder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
;  TITLE OF INVENTION:  TRANSPORT
;  FILE REFERENCE: BGI-125CPCN
;  CURRENT APPLICATION NUMBER: US/10/627,476
;  CURRENT FILING DATE:  2003-07-25
;  PRIOR APPLICATION NUMBER: 09/602,787
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: USSN 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: DE 19931454.3
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931478.0
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931563.9
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932122.1
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932124.8
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932125.6
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932128.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932180.9
;  PRIOR FILING DATE: 1999-07-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 206
;   LENGTH: 438
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-10-627-476-206

  Query Match             53.6%;  Score 45;  DB 4;  Length 438;
  Best Local Similarity   62.5%;  Pred. No. 2.2e+02;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              |||||   | || | |
Db        419 KSYGGAYCVMGSKDMG 434


RESULT 15
US-10-282-122A-52320
; Sequence 52320, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
```

```
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52320
;   LENGTH: 513
;   TYPE: PRT
;   ORGANISM: Clostridium botulinum
US-10-282-122A-52320

  Query Match             53.6%;  Score 45;  DB 4;  Length 513;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+02;
  Matches    9;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              ||| |||||  |
Db        232 YGGIEVVGGLVD 243


Search completed: March 21, 2006, 21:52:52
Job time : 42.8909 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:50:33 ; Search time 6.98182 Seconds
(without alignments)
65.594 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	50.0	501	7	US-11-167-856-22	Sequence 22, Appl
2	42	50.0	548	6	US-10-858-730-68	Sequence 68, Appl
3	42	50.0	910	7	US-11-206-587-17	Sequence 17, Appl
4	41	48.8	449	6	US-10-878-556A-57	Sequence 57, Appl
5	41	48.8	449	6	US-10-500-941-5	Sequence 5, Appli
6	41	48.8	571	7	US-11-121-438-12	Sequence 12, Appl
7	41	48.8	845	7	US-11-096-568A-28842	Sequence 28842, A
8	41	48.8	897	7	US-11-096-568A-28841	Sequence 28841, A
9	41	48.8	912	7	US-11-096-568A-28840	Sequence 28840, A
10	40	47.6	329	7	US-11-096-568A-23511	Sequence 23511, A
11	40	47.6	352	7	US-11-096-568A-23510	Sequence 23510, A
12	40	47.6	397	6	US-10-858-730-29	Sequence 29, Appl
13	40	47.6	413	6	US-10-880-764-3	Sequence 3, Appli
14	40	47.6	424	7	US-11-240-769-64	Sequence 64, Appl
15	40	47.6	499	7	US-11-087-099-11253	Sequence 11253, A
16	40	47.6	772	6	US-10-330-773-67	Sequence 67, Appl
17	39	46.4	291	7	US-11-087-099-3763	Sequence 3763, Ap
18	39	46.4	616	7	US-11-096-568A-1435	Sequence 1435, Ap
19	39	46.4	646	7	US-11-096-568A-1434	Sequence 1434, Ap
20	39	46.4	754	6	US-10-330-773-72	Sequence 72, Appl
21	39	46.4	813	7	US-11-096-568A-31851	Sequence 31851, A
22	39	46.4	851	7	US-11-096-568A-31850	Sequence 31850, A
23	39	46.4	902	7	US-11-096-568A-31849	Sequence 31849, A
24	38.5	45.8	178	7	US-11-087-099-4090	Sequence 4090, Ap
25	38.5	45.8	185	7	US-11-087-099-9737	Sequence 9737, Ap
26	38.5	45.8	185	7	US-11-087-099-9811	Sequence 9811, Ap
27	38	45.2	126	7	US-11-087-099-10637	Sequence 10637, A
28	38	45.2	201	7	US-11-096-568A-6894	Sequence 6894, Ap
29	38	45.2	257	7	US-11-096-568A-3960	Sequence 3960, Ap
30	38	45.2	260	7	US-11-096-568A-6893	Sequence 6893, Ap
31	38	45.2	304	6	US-10-793-626-2190	Sequence 2190, Ap
32	38	45.2	306	7	US-11-096-568A-3959	Sequence 3959, Ap
33	38	45.2	363	7	US-11-087-099-1891	Sequence 1891, Ap
34	38	45.2	363	7	US-11-087-099-6271	Sequence 6271, Ap
35	38	45.2	363	7	US-11-087-099-6638	Sequence 6638, Ap
36	38	45.2	363	7	US-11-087-099-9995	Sequence 9995, Ap
37	38	45.2	363	7	US-11-087-099-10918	Sequence 10918, A
38	38	45.2	364	7	US-11-087-099-9837	Sequence 9837, Ap
39	38	45.2	365	7	US-11-087-099-10907	Sequence 10907, A
40	38	45.2	366	7	US-11-087-099-1520	Sequence 1520, Ap
41	38	45.2	369	7	US-11-096-568A-3958	Sequence 3958, Ap
42	38	45.2	436	7	US-11-087-099-383	Sequence 383, App
43	38	45.2	447	7	US-11-087-099-8364	Sequence 8364, Ap
44	38	45.2	484	7	US-11-052-554A-170	Sequence 170, App
45	38	45.2	491	7	US-11-024-959-363	Sequence 363, App

ALIGNMENTS

RESULT 1
US-11-167-856-22
; Sequence 22, Application US/11167856
; Publication No. US20050268352A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurtele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; TITLE OF INVENTION: Acetyl CoA Levels in Plants
; FILE REFERENCE: P2194USDIV-2
; CURRENT APPLICATION NUMBER: US/11/167,856
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/293,865
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 22
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-11-167-856-22

Query Match 50.0%; Score 42; DB 7; Length 501;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 YGGYEVVGGSSDSG 16
              ||||:: |   :||
Db        465 YGGYKMSGNCRESG 478
```

RESULT 2
US-10-858-730-68
; Sequence 68, Application US/10858730
; Publication No. US20050255568A1

```
; GENERAL INFORMATION:
;  APPLICANT: Bailey, Richard B.
;  APPLICANT:  Blomquist, Paul
;  APPLICANT:  Doten, Reed
;  APPLICANT:  Driggers, Edward M.
;  APPLICANT:  Madden, Kevin T.
;  APPLICANT:  O'Leary, Jessica
;  APPLICANT:  O'Toole, George
;  APPLICANT:  Trueheart, Joshua
;  APPLICANT:  Walbridge, Michael J.
;  APPLICANT:  Yorgey, Peter S.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
;  TITLE OF INVENTION:  PRODUCTION
;  FILE REFERENCE: 14184-030001
;  CURRENT APPLICATION NUMBER: US/10/858,730
;  CURRENT FILING DATE:  2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/475,000
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: US 60/551,860
;  PRIOR FILING DATE: 2004-03-10
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
;   LENGTH: 548
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-10-858-730-68

  Query Match             50.0%;  Score 42;  DB 6;  Length 548;
  Best Local Similarity   56.2%;  Pred. No. 62;
  Matches    9;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              |:|||   | || | |
Db        430 KAYGGAYCVMGSKDMG 445


RESULT 3
US-11-206-587-17
; Sequence 17, Application US/11206587
; Publication No. US20060040306A1
; GENERAL INFORMATION:
;  APPLICANT: Leiby, Kevin R.
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Glucksmann, Maria A.
;  TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
;  TITLE OF INVENTION:  14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 381552004900
;  CURRENT APPLICATION NUMBER: US/11/206,587
;  CURRENT FILING DATE:  2005-08-18
;  PRIOR APPLICATION NUMBER: US/10/012,140
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 60/246,768
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/246,772
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/249,185
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 910
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-206-587-17

  Query Match             50.0%;  Score 42;  DB 7;  Length 910;
  Best Local Similarity   81.8%;  Pred. No. 1e+02;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 GYEVVGGSSDS 15
              |||||| || |
Db        211 GYEVVGSSSAS 221


RESULT 4
US-10-878-556A-57
; Sequence 57, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
;  APPLICANT: Hoffmann La-Roche Inc.
;  TITLE OF INVENTION: HCV regulated protein expression
;  FILE REFERENCE: 21762
;  CURRENT APPLICATION NUMBER: US/10/878,556A
;  CURRENT FILING DATE:  2004-06-28
;  NUMBER OF SEQ ID NOS: 199
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: sw_hum/roh2_human
;   DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-57

  Query Match             48.8%;  Score 41;  DB 6;  Length 449;
  Best Local Similarity   58.3%;  Pred. No. 71;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              ||||:  || :|
Db        240 YGGYDDYGGYND 251


RESULT 5
US-10-500-941-5
; Sequence 5, Application US/10500941
; Publication No. US20060052293A1
; GENERAL INFORMATION:
;  APPLICANT: Climent-Johansson, Isabel
;  APPLICANT:  Enerback, Sven
;  TITLE OF INVENTION: PROTEIN COMPLEXES
;  FILE REFERENCE: 13425-102US1
;  CURRENT APPLICATION NUMBER: US/10/500,941
;  CURRENT FILING DATE:  2004-07-08
;  PRIOR APPLICATION NUMBER: PCT/SE03/00139
;  PRIOR FILING DATE: 2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/377,349
;  PRIOR FILING DATE: 2002-04-30
;  PRIOR APPLICATION NUMBER: SE 0200265-7
;  PRIOR FILING DATE: 2002-01-29
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-500-941-5

  Query Match             48.8%;  Score 41;  DB 6;  Length 449;
  Best Local Similarity   58.3%;  Pred. No. 71;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              ||||:  || :|
Db        240 YGGYDDYGGYND 251


RESULT 6
US-11-121-438-12
; Sequence 12, Application US/11121438
; Publication No. US20060014173A1
; GENERAL INFORMATION:
;  APPLICANT: Huang, Shi
```

```
;  TITLE OF INVENTION: PR/SET- Domain Containing Nucleic Acids,
;  TITLE OF INVENTION:  Polypeptides, Antibodies and Methods of Use
;  FILE REFERENCE: P-LJ 5301
;  CURRENT APPLICATION NUMBER: US/11/121,438
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: US/10/200,012
;  PRIOR FILING DATE: 2002-07-18
;  PRIOR APPLICATION NUMBER: US 09/910,478
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 571
;   TYPE: PRT
;   ORGANISM: Human
US-11-121-438-12

  Query Match             48.8%;  Score 41;  DB 7;  Length 571;
  Best Local Similarity   41.7%;  Pred. No. 91;
  Matches    5;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 GYEVVGGSSDSG 16
              |::::|| ::||
Db        130 GHQIIGGDNESG 141


RESULT 7
US-11-096-568A-28842
; Sequence 28842, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28842
;   LENGTH: 845
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(845)
;   OTHER INFORMATION: Ceres Seq. ID no. 3081942
US-11-096-568A-28842

  Query Match             48.8%;  Score 41;  DB 7;  Length 845;
  Best Local Similarity   56.2%;  Pred. No. 1.4e+02;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              :| ||  | |||| :|
Db        676 ESLGGSSVYGGSSVNG 691


RESULT 8
US-11-096-568A-28841
; Sequence 28841, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28841
;   LENGTH: 897
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(897)
;   OTHER INFORMATION: Ceres Seq. ID no. 3081941
US-11-096-568A-28841

  Query Match             48.8%;  Score 41;  DB 7;  Length 897;
  Best Local Similarity   56.2%;  Pred. No. 1.4e+02;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              :| ||  | |||| :|
Db        728 ESLGGSSVYGGSSVNG 743


RESULT 9
US-11-096-568A-28840
; Sequence 28840, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28840
;   LENGTH: 912
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(912)
;   OTHER INFORMATION: Ceres Seq. ID no. 3081940
US-11-096-568A-28840

  Query Match             48.8%;  Score 41;  DB 7;  Length 912;
  Best Local Similarity   56.2%;  Pred. No. 1.5e+02;
  Matches    9;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              :| ||  | |||| :|
Db        743 ESLGGSSVYGGSSVNG 758


RESULT 10
US-11-096-568A-23511
; Sequence 23511, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23511
;   LENGTH: 329
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(329)
;   OTHER INFORMATION: Ceres Seq. ID no. 12413226
US-11-096-568A-23511

  Query Match             47.6%;  Score 40;  DB 7;  Length 329;
  Best Local Similarity   53.8%;  Pred. No. 74;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDSG 16
```

```
              |||:  ||: | |
Db        312 GGYDDYGGNDDYG 324


RESULT 11
US-11-096-568A-23510
; Sequence 23510, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23510
;   LENGTH: 352
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(352)
;   OTHER INFORMATION: Ceres Seq. ID no. 12413225
US-11-096-568A-23510

  Query Match             47.6%;  Score 40;  DB 7;  Length 352;
  Best Local Similarity   53.8%;  Pred. No. 79;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDSG 16
              |||:  ||: | |
Db        335 GGYDDYGGNDDYG 347


RESULT 12
US-10-858-730-29
; Sequence 29, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
;  APPLICANT: Bailey, Richard B.
;  APPLICANT:  Blomquist, Paul
;  APPLICANT:  Doten, Reed
;  APPLICANT:  Driggers, Edward M.
;  APPLICANT:  Madden, Kevin T.
;  APPLICANT:  O'Leary, Jessica
;  APPLICANT:  O'Toole, George
;  APPLICANT:  Trueheart, Joshua
;  APPLICANT:  Walbridge, Michael J.
;  APPLICANT:  Yorgey, Peter S.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
;  TITLE OF INVENTION:  PRODUCTION
;  FILE REFERENCE: 14184-030001
;  CURRENT APPLICATION NUMBER: US/10/858,730
;  CURRENT FILING DATE:  2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/475,000
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: US 60/551,860
;  PRIOR FILING DATE: 2004-03-10
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 397
;   TYPE: PRT
;   ORGANISM: Thermobifida fusca
US-10-858-730-29

  Query Match             47.6%;  Score 40;  DB 6;  Length 397;
  Best Local Similarity   53.3%;  Pred. No. 89;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDSG 16
              :||||   || : ||
Db        263 TYGGYARHGGGAFSG 277


RESULT 13
US-10-880-764-3
; Sequence 3, Application US/10880764
; Publication No. US20060003335A1
; GENERAL INFORMATION:
;  APPLICANT: John D. Crispino
;  TITLE OF INVENTION: METHODS FOR DIAGNOSING ACUTE MEGAKARYOBLASTIC LEUKEMIA
;  FILE REFERENCE: 27373/38893A
;  CURRENT APPLICATION NUMBER: US/10/880,764
;  CURRENT FILING DATE:  2004-06-30
;  PRIOR APPLICATION NUMBER: US 60/483,954
;  PRIOR FILING DATE: 2003-07-01
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 413
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-880-764-3

  Query Match             47.6%;  Score 40;  DB 6;  Length 413;
  Best Local Similarity   61.5%;  Pred. No. 93;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDSG 16
              ||: || | | ||
Db        332 GGFMVVAGGSGSG 344


RESULT 14
US-11-240-769-64
; Sequence 64, Application US/11240769
; Publication No. US20060036089A1
; GENERAL INFORMATION:
;  APPLICANT: Soppet et al.
;  TITLE OF INVENTION: 33 Human Secreted Proteins
;  FILE REFERENCE: PZ037P1C2
;  CURRENT APPLICATION NUMBER: US/11/240,769
;  CURRENT FILING DATE:  2005-10-03
;  PRIOR APPLICATION NUMBER: 09/997,131
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 09/628,508
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: PCT/US00/03062
;  PRIOR FILING DATE: 2000-02-08
;  PRIOR APPLICATION NUMBER: 60/119,468
;  PRIOR FILING DATE: 1999-02-10
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
;   LENGTH: 424
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (268)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (316)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (318)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-240-769-64

  Query Match             47.6%;  Score 40;  DB 7;  Length 424;
  Best Local Similarity   53.8%;  Pred. No. 95;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          2 SYGGYEVVGGSSD 14
              :|| :| | ||:|
Db         62 NYGSFENVSGSTD 74


RESULT 15
US-11-087-099-11253
; Sequence 11253, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11253
;   LENGTH: 499
;   TYPE: PRT
;   ORGANISM: Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(499)
;   OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-11253

  Query Match             47.6%;  Score 40;  DB 7;  Length 499;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSDSG 16
              | ||  :| |:| |
Db        110 YDGYSFIGLSTDGG 123


Search completed: March 21, 2006, 21:53:22
Job time : 6.98182 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:04 ; Search time 14.8364 Seconds
(without alignments)
89.160 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	92.9	529	2	US-09-381-656-1	Sequence 1, Appli
2	78	92.9	529	2	US-10-164-230-61	Sequence 61, Appl
3	46	54.8	289	1	US-07-593-657-7	Sequence 7, Appli
4	46	54.8	289	2	US-08-942-012B-4	Sequence 4, Appli
5	45	53.6	143	2	US-09-252-991A-24446	Sequence 24446, A
6	45	53.6	173	2	US-09-602-787A-210	Sequence 210, App
7	45	53.6	291	1	US-07-593-657-5	Sequence 5, Appli
8	45	53.6	291	2	US-09-029-213B-2	Sequence 2, Appli
9	45	53.6	291	2	US-09-029-213B-4	Sequence 4, Appli
10	45	53.6	291	2	US-08-942-012B-2	Sequence 2, Appli
11	45	53.6	438	2	US-09-602-787A-206	Sequence 206, App
12	45	53.6	894	2	US-09-252-991A-21126	Sequence 21126, A
13	43	51.2	361	2	US-09-902-540-10088	Sequence 10088, A
14	42	50.0	414	2	US-09-949-016-8543	Sequence 8543, Ap
15	42	50.0	415	2	US-09-949-016-6571	Sequence 6571, Ap
16	42	50.0	465	2	US-09-252-991A-29387	Sequence 29387, A
17	42	50.0	501	2	US-09-344-882-22	Sequence 22, Appl
18	42	50.0	501	2	US-10-293-865-22	Sequence 22, Appl
19	42	50.0	507	1	US-08-363-475-19	Sequence 19, Appl
20	42	50.0	515	2	US-09-489-039A-8402	Sequence 8402, Ap
21	42	50.0	532	1	US-08-363-475-22	Sequence 22, Appl
22	42	50.0	643	2	US-08-797-358B-3	Sequence 3, Appli
23	42	50.0	691	2	US-09-902-540-16476	Sequence 16476, A
24	42	50.0	889	2	US-09-902-540-15702	Sequence 15702, A
25	42	50.0	1045	1	US-08-452-083-2	Sequence 2, Appli
26	41.5	49.4	675	2	US-09-134-001C-4547	Sequence 4547, Ap
27	41	48.8	68	2	US-09-518-550-32	Sequence 32, Appl
28	41	48.8	154	2	US-09-252-991A-26700	Sequence 26700, A
29	41	48.8	308	2	US-09-513-999C-4832	Sequence 4832, Ap
30	41	48.8	396	2	US-09-273-686-2	Sequence 2, Appli
31	41	48.8	396	2	US-09-583-110-2778	Sequence 2778, Ap
32	41	48.8	405	2	US-09-107-433-2830	Sequence 2830, Ap
33	41	48.8	460	2	US-09-917-254-79	Sequence 79, Appl
34	41	48.8	571	2	US-10-200-012-12	Sequence 12, Appl
35	41	48.8	901	2	US-09-134-001C-5351	Sequence 5351, Ap
36	40.5	48.2	373	2	US-09-949-016-10485	Sequence 10485, A
37	40	47.6	23	2	US-09-045-632-59	Sequence 59, Appl
38	40	47.6	331	2	US-09-949-016-7015	Sequence 7015, Ap
39	40	47.6	401	2	US-09-540-236-3161	Sequence 3161, Ap
40	40	47.6	404	2	US-09-107-532A-6821	Sequence 6821, Ap
41	40	47.6	449	2	US-09-252-991A-20941	Sequence 20941, A
42	40	47.6	559	2	US-09-583-110-3735	Sequence 3735, Ap
43	40	47.6	561	2	US-09-252-991A-28074	Sequence 28074, A
44	40	47.6	567	2	US-09-107-433-4592	Sequence 4592, Ap
45	40	47.6	762	2	US-09-724-653-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Appl.

Query Match 92.9%; Score 78; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSYGGYEVVGGSSDS 15
||||||||||||||||
Db 291 KSYGGYEVVGGSSDS 305

RESULT 2
US-10-164-230-61
; Sequence 61, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             92.9%;  Score 78;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305


RESULT 3
US-07-593-657-7
; Sequence 7, Application US/07593657
; Patent No. 5266317
;  GENERAL INFORMATION:
;    APPLICANT:  Tomalski, Michael D.
;    APPLICANT:  Miller, Lois K.
;    TITLE OF INVENTION:  INSECT-SPECIFIC PARALYTIC NEUROTOXIN
;    TITLE OF INVENTION:  GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  CO
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/593,657
;      FILING DATE:  19901004
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  14-90
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-593-657-7

  Query Match             54.8%;  Score 46;  DB 1;  Length 289;
  Best Local Similarity   46.2%;  Pred. No. 27;
  Matches    6;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| :
Db         80 KVYGGFDIIGGKA 92


RESULT 4
US-08-942-012B-4
; Sequence 4, Application US/08942012B
; Patent No. 6235278
; GENERAL INFORMATION:
;  APPLICANT:  Miller, Lois K.
;  APPLICANT:  Lu, Albert
;  APPLICANT:  Dierks, Peter
;  APPLICANT:  Black, Bruce
;  TITLE OF INVENTION:  Biological Insect Control  Agents Expressing
;  TITLE OF INVENTION:  Insect-Specific Toxin Genes, Methods and Compositions
;  FILE REFERENCE:  28-96a
;  CURRENT APPLICATION NUMBER: US/08/942,012B
;  CURRENT FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  08/729,606
;  PRIOR FILING DATE:  2000-10-01
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Pyemotes tritici
US-08-942-012B-4

  Query Match             54.8%;  Score 46;  DB 2;  Length 289;
  Best Local Similarity   46.2%;  Pred. No. 27;
  Matches    6;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| :
Db         80 KVYGGFDIIGGKA 92


RESULT 5
US-09-252-991A-24446
```

```
; Sequence 24446, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 24446
;   LENGTH: 143
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24446

  Query Match             53.6%;  Score 45;  DB 2;  Length 143;
  Best Local Similarity   69.2%;  Pred. No. 18;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSD 14
              ||  ||||| ||:
Db         74 SYSRYEVVGSSSE 86


RESULT 6
US-09-602-787A-210
; Sequence 210, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Mark
;  APPLICANT:  Kr"ger, Burkhard
;  APPLICANT:  Sch"der, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
;  TITLE OF INVENTION:  TRANSPORT
;  FILE REFERENCE: BGI-125CP
;  CURRENT APPLICATION NUMBER: US/09/602,787A
;  CURRENT FILING DATE:  2000-06-23
;  PRIOR APPLICATION NUMBER: USSN 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: DE 19931454.3
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931478.0
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19931563.9
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932122.1
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932124.8
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932125.6
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932128.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932180.9
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932182.5
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932190.6
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932191.4
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932209.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932212.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932227.9
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932228.7
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932229.5
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932230.9
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932927.3
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: DE 19933005.0
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: DE 19933006.9
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: DE 19940764.9
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940765.7
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940766.5
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940830.0
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940831.9
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940832.7
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19940833.5
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: DE 19941378.9
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: DE 19941379.7
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: DE 19941395.9
;  PRIOR FILING DATE: 1999-08-31
;  PRIOR APPLICATION NUMBER: DE 19942077.7
;  PRIOR FILING DATE: 1999-09-03
;  PRIOR APPLICATION NUMBER: DE 19942078.5
;  PRIOR FILING DATE: 1999-09-03
;  PRIOR APPLICATION NUMBER: DE 19942079.3
;  PRIOR FILING DATE: 1999-09-03
;  PRIOR APPLICATION NUMBER: DE 19942088.2
;  PRIOR FILING DATE: 1999-09-03
;  NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 210
;   LENGTH: 173
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 70
;   OTHER INFORMATION: Xaa indicates any amino acid
;   NAME/KEY: unsure
;   LOCATION: 72
;   OTHER INFORMATION: Xaa indicates any amino acid
;   NAME/KEY: unsure
;   LOCATION: 86
;   OTHER INFORMATION: Xaa indicates any amino acid
US-09-602-787A-210

  Query Match             53.6%;  Score 45;  DB 2;  Length 173;
  Best Local Similarity   62.5%;  Pred. No. 22;
  Matches   10;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              |||||   | || | |
Db        154 KSYGGAYCVMGSKDMG 169


RESULT 7
US-07-593-657-5
; Sequence 5, Application US/07593657
; Patent No. 5266317
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Tomalski, Michael D.
;    APPLICANT:  Miller, Lois K.
;    TITLE OF INVENTION:  INSECT-SPECIFIC PARALYTIC NEUROTOXIN
;    TITLE OF INVENTION:  GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
;    TITLE OF INVENTION:  COMPOSITIONS
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee and Winner, P.C.
;      STREET:  5370 Manhattan Circle, Suite 201
;      CITY:  Boulder
;      STATE:  CO
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/593,657
;      FILING DATE:  19901004
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  14-90
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  291 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-593-657-5

  Query Match             53.6%;  Score 45;  DB 1;  Length 291;
  Best Local Similarity   54.5%;  Pred. No. 38;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              | |||::::||
Db         87 KVYGGFDIIGG 97


RESULT 8
US-09-029-213B-2
; Sequence 2, Application US/09029213B
; Patent No. 6180098
;  GENERAL INFORMATION:
;    APPLICANT:  CHRISTIAN, Peter D.
;    TITLE OF INVENTION:  RECOMBINANT HELICOVERPA BACULOVIRUSES
;    TITLE OF INVENTION:  EXPRESSING HETEROLOGOUS DNA
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDermott, Will & Emery
;      STREET:  600 13th Street, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,213B
;      FILING DATE:  31-AUG-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Joseph Hyosuk Kim
;      REGISTRATION NUMBER:  41,425
;      REFERENCE/DOCKET NUMBER:  50179-048
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-756-8000
;      TELEFAX:  202-756-8087
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  291 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-029-213B-2

  Query Match             53.6%;  Score 45;  DB 2;  Length 291;
  Best Local Similarity   54.5%;  Pred. No. 38;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              | |||::::||
Db         87 KVYGGFDIIGG 97


RESULT 9
US-09-029-213B-4
; Sequence 4, Application US/09029213B
; Patent No. 6180098
;  GENERAL INFORMATION:
;    APPLICANT:  CHRISTIAN, Peter D.
;    TITLE OF INVENTION:  RECOMBINANT HELICOVERPA BACULOVIRUSES
;    TITLE OF INVENTION:  EXPRESSING HETEROLOGOUS DNA
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDermott, Will & Emery
;      STREET:  600 13th Street, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,213B
;      FILING DATE:  31-AUG-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Joseph Hyosuk Kim
;      REGISTRATION NUMBER:  41,425
;      REFERENCE/DOCKET NUMBER:  50179-048
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-756-8000
;      TELEFAX:  202-756-8087
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  291 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-029-213B-4

  Query Match             53.6%;  Score 45;  DB 2;  Length 291;
  Best Local Similarity   54.5%;  Pred. No. 38;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              | |||::::||
Db         87 KVYGGFDIIGG 97


RESULT 10
US-08-942-012B-2
```

; Sequence 2, Application US/08942012B
; Patent No. 6235278
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Lu, Albert
; APPLICANT: Dierks, Peter
; APPLICANT: Black, Bruce
; TITLE OF INVENTION: Biological Insect Control Agents Expressing
; TITLE OF INVENTION: Insect-Specific Toxin Genes, Methods and Compositions
; FILE REFERENCE: 28-96a
; CURRENT APPLICATION NUMBER: US/08/942,012B
; CURRENT FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: 08/729,606
; PRIOR FILING DATE: 2000-10-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Pyemotes tritici
US-08-942-012B-2

Query Match 53.6%; Score 45; DB 2; Length 291;
Best Local Similarity 54.5%; Pred. No. 38;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGG 11
              | |||:::::||
Db         87 KVYGGFDIIGG 97
```

RESULT 11
US-09-602-787A-206
; Sequence 206, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 206
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-206

Query Match 53.6%; Score 45; DB 2; Length 438;
Best Local Similarity 62.5%; Pred. No. 58;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSSDSG 16
              |||||   | || | |
Db        419 KSYGGAYCVMGSKDMG 434
```

RESULT 12
US-09-252-991A-21126
; Sequence 21126, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

```
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 21126
;   LENGTH: 894
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21126

  Query Match             53.6%;  Score 45;  DB 2;  Length 894;
  Best Local Similarity   43.8%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              : || :|:|||    |
Db        114 RQYGQFEIVGGEQGEG 129


RESULT 13
US-09-902-540-10088
; Sequence 10088, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10088
;   LENGTH: 361
;   TYPE: PRT
;   ORGANISM: Myxococcus xanthus
US-09-902-540-10088

  Query Match             51.2%;  Score 43;  DB 2;  Length 361;
  Best Local Similarity   77.8%;  Pred. No. 94;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGG 11
              |||| :|||
Db        100 YGGYGIVGG 108


RESULT 14
US-09-949-016-8543
; Sequence 8543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8543
;   LENGTH: 414
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8543

  Query Match             50.0%;  Score 42;  DB 2;  Length 414;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              |||||   | ||
Db        239 YGGYEEYSGLSD 250


RESULT 15
US-09-949-016-6571
; Sequence 6571, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6571
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6571

  Query Match             50.0%;  Score 42;  DB 2;  Length 415;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGGYEVVGGSSD 14
              |||||   | ||
Db        240 YGGYEEYSGLSD 251


Search completed: March 21, 2006, 21:50:22
Job time : 15.8364 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:46:13 ; Search time 6.72727 Seconds
(without alignments)
143.025 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	502	2	A48679	differentiated ker
2	42	75.0	644	2	T33730	hypothetical prote
3	41	73.2	281	2	H84720	probable endonucle
4	41	73.2	402	2	A86194	hypothetical prote
5	39	69.6	332	2	C95150	SAP domain protein
6	39	69.6	424	2	B56144	aspartate carbamoy
7	39	69.6	561	2	A87627	PDZ domain family
8	39	69.6	823	2	H83724	hypothetical prote
9	39	69.6	1330	2	A36373	hypothetical prote
10	39	69.6	1333	2	T38401	retrotransposable
11	37	66.1	142	2	A90432	hypothetical prote
12	37	66.1	153	2	A81110	hypothetical prote
13	37	66.1	282	2	G70482	pantothenate synth
14	37	66.1	356	2	D84280	hypothetical prote
15	37	66.1	361	2	AF1593	B. subtilis PBSX p
16	37	66.1	468	2	AF2163	hypothetical prote
17	37	66.1	1033	2	T49179	Ca2+-transporting
18	37	66.1	1082	2	T15269	hypothetical prote
19	36.5	65.2	728	2	S59964	procollagen-lysine
20	36.5	65.2	730	2	A23742	procollagen-lysine
21	36	64.3	65	2	S42675	adhesive protein -
22	36	64.3	136	1	D71099	hypothetical prote
23	36	64.3	246	2	H72497	probable NH(3)-dep
24	36	64.3	248	2	D64242	sporulation protei
25	36	64.3	248	2	S73606	probable NH(3)-dep
26	36	64.3	276	2	T52316	chlorophyll a/b-bi
27	36	64.3	423	2	G83595	noncatalytic dihyd
28	36	64.3	546	2	B37491	major capsid prote
29	36	64.3	664	2	T28852	probable potassium
30	35	62.5	177	2	D95278	conserved hypothet
31	35	62.5	245	1	NDECR5	type II site-speci
32	35	62.5	277	2	H86589	hypothetical prote
33	35	62.5	277	2	C72033	hypothetical prote
34	35	62.5	363	2	C69255	first mannosyl tra
35	35	62.5	402	2	S23774	triose phosphate/3
36	35	62.5	410	2	T06213	probable aspartic
37	35	62.5	410	2	T04372	protein EEA1 - bar
38	35	62.5	707	2	T00665	hypothetical prote
39	35	62.5	1555	2	S38758	amylo-alpha-1,6-gl
40	34	60.7	75	2	S72288	ribosomal protein
41	34	60.7	84	2	E97050	hypothetical prote
42	34	60.7	140	2	E83391	probable ring-clea
43	34	60.7	209	2	C95167	hypothetical prote
44	34	60.7	213	2	H70325	endonuclease III -
45	34	60.7	260	2	AC3456	hydrolases of the

ALIGNMENTS

RESULT 1
A48679
differentiated keratinocyte S protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A48679
R;Zhou, Y.; Chaplin, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 9470-9474, 1993
A;Title: Identification in the HLA class I region of a gene expressed late in keratinoc
A;Reference number: A48679; MUID:94022396; PMID:8415725
A;Accession: A48679
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <ZHO>
A;Cross-references: UNIPROT:O43509; UNIPARC:UPI0000177376; GB:L20815
A;Note: sequence shown follows the authors' translation at positions 47 and 410
C;Genetics:
A;Map position: 6p21.3
C;Superfamily: loricrin

Query Match 100.0%; Score 56; DB 2; Length 502;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326
```

RESULT 2
T33730
hypothetical protein - Zymomonas mobilis
C;Species: Zymomonas mobilis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 28-Jul-2003
C;Accession: T33730
R;Lee, J.S.; Jin, S.J.; Kang, H.L.; Kang, H.S.
submitted to the EMBL Data Library, August 1998
A;Description: Sequence analysis of 67E10 cosmid clone of Zymomonas mobilis ZM4.
A;Reference number: Z21392
A;Accession: T33730
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-644 <LEE>
A;Cross-references: UNIPARC:UPI00000B8186; EMBL:AF086791; NID:g3820581; PID:g3820589; P
C;Superfamily: glycyl monoaminopeptidase

Query Match 75.0%; Score 42; DB 2; Length 644;
Best Local Similarity 66.7%; Pred. No. 4.4;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 IYPVGYFTK 10
              :|| ||||:
```

```
Db        165 LYPAGYFTR 173


RESULT 3
H84720
probable endonuclease [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84720
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84720
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <STO>
A;Cross-references: UNIPROT:Q9SIC4; UNIPARC:UPI000017A6CE; GB:AE002093; NID:g4589956; PI
C;Genetics:
A;Gene: At2g31450
A;Map position: 2

  Query Match             73.2%;  Score 41;  DB 2;  Length 281;
  Best Local Similarity   66.7%;  Pred. No. 3;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              |||||::|:
Db        123 IYPVGFYTR 131


RESULT 4
A86194
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86194
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A86194
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <STO>
A;Cross-references: UNIPROT:Q9MA35; UNIPARC:UPI00000A87F1; GB:AE005172; NID:g6850320; PI
C;Genetics:
A;Map position: 1

  Query Match             73.2%;  Score 41;  DB 2;  Length 402;
  Best Local Similarity   66.7%;  Pred. No. 4.3;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              |||||::|:
Db        228 IYPVGFYTR 236


RESULT 5
C95150
SAP domain protein [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95150
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: C95150
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <KUR>
A;Cross-references: UNIPROT:Q97QC7; UNIPARC:UPI00000517B2; GB:AE005672; PIDN:AAK75396.1,
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1292

  Query Match             69.6%;  Score 39;  DB 2;  Length 332;
  Best Local Similarity   60.0%;  Pred. No. 8.6;
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ::| :|||||
Db         73 QLYKLGYFTK 82


RESULT 6
B56144
aspartate carbamoyltransferase (EC 2.1.3.2) 44k noncatalytic chain - Pseudomonas putida
C;Species: Pseudomonas putida
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: B56144
R;Schurr, M.J.; Vickrey, J.F.; Kumar, A.P.; Campbell, A.L.; Cunin, R.; Benjamin, R.C.; &
J. Bacteriol. 177, 1751-1759, 1995
A;Title: Aspartate transcarbamoylase genes of Pseudomonas putida: requirement for an ina
A;Reference number: A56144; MUID:95204340; PMID:7896697
A;Accession: B56144
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-424 <SCH>
A;Cross-references: UNIPROT:Q59712; UNIPARC:UPI0000132E1F; GB:M97254; NID:g151523; PIDN
C;Genetics:
A;Gene: pyrC'
C;Superfamily: Bacillus dihydroorotase; Bacillus dihydroorotase homology
C;Keywords: transferase
F;47-424/Domain: Bacillus dihydroorotase homology <DHO>

  Query Match             69.6%;  Score 39;  DB 2;  Length 424;
  Best Local Similarity   60.0%;  Pred. No. 11;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              |:||:|  ||
Db        120 KVYPIGALTK 129


RESULT 7
A87627
PDZ domain family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 28-Jul-2003
C;Accession: A87627
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-561 <STO>
A;Cross-references: UNIPARC:UPI00000C792F; GB:AE005673; NID:g13424697; PIDN:AAK25013.1;
C;Genetics:
```

A;Gene: CC3051
C;Superfamily: glycyl monoaminopeptidase

Query Match 69.6%; Score 39; DB 2; Length 561;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IYPVGYFTK 10
:|| ||:|:
Db 82 LYPAGYYTR 90

RESULT 8
H83724
hypothetical protein BH0600 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83724
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-823 <STO>
A;Cross-references: UNIPROT:Q9KF85; UNIPARC:UPI00000C391F; GB:AP001509; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0600

Query Match 69.6%; Score 39; DB 2; Length 823;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IYPVGYFT 9
:||||:||
Db 392 LYPVGHFT 399

RESULT 9
A36373
hypothetical protein Tf1 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 09-Jul-2004
C;Accession: A36373
R;Levin, H.L.; Weaver, D.C.; Boeke, J.D.
Mol. Cell. Biol. 10, 6791-6798, 1990
A;Title: Two related families of retrotransposons from Schizosaccharomyces pombe.
A;Reference number: A36373; MUID:91061791; PMID:2174117
A;Accession: A36373
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1330 <LEV>
A;Cross-references: UNIPROT:Q01910; UNIPARC:UPI000006B888; GB:M38526; NID:g173476; PIDN:

Query Match 69.6%; Score 39; DB 2; Length 1330;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYPVGYFT 9
| |||||::
Db 728 KYYPVGYYS 736

RESULT 10
T38401
retrotransposable element Tf2 155K protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38401; T39189; T40783; T37743; JN0791; T38456; T40836; T39868; T50374; T38
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21791
A;Accession: T38401
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1333 <MCL>
A;Cross-references: UNIPROT:Q05654; UNIPARC:UPI0000135298; EMBL:Z69240; PIDN:CAA93236.1
A;Experimental source: strain 972h-; cosmid 26A3
A;Genetics: MCL1
R;Wedler, H.; Duesterhoeft, A.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21834
A;Accession: T39189
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1333 <WED>
A;Cross-references: UNIPARC:UPI0000135298; EMBL:AL121764; PIDN:CAB57422.1; GSPDB:GN00066
A;Experimental source: strain 972h-; cosmid c9
A;Genetics: WED1
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21875
A;Accession: T40783
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1333 <WOO>
A;Cross-references: UNIPARC:UPI0000135298; EMBL:AL049769; PIDN:CAB42363.1; GSPDB:GN0006
A;Experimental source: strain 972h-; cosmid c9B6
R;Rieger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A;Reference number: Z21743
A;Accession: T37743
A;Molecule type: DNA
A;Residues: 120-1333 <RIE>
A;Cross-references: UNIPARC:UPI000016203E; EMBL:AL035248; PIDN:CAB58169.1; GSPDB:GN00066
R;Weaver, D.C.; Shpakovski, G.V.; Caputo, E.; Levin, H.L.; Boeke, J.D.
Gene 131, 135-139, 1993
A;Title: Sequence analysis of closely related retrotransposon families from fission yeas
A;Reference number: JN0791; MUID:93380663; PMID:8396547
A;Accession: JN0791
A;Molecule type: DNA
A;Residues: 1-205,'P',207-1333 <WEA>
A;Cross-references: UNIPARC:UPI0000161F17; GB:L10324; NID:g173438; PIDN:AAA91215.1; PID
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21749
A;Accession: T38456
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-205,'P',207-1333 <MUR>
A;Cross-references: UNIPARC:UPI0000161F17; EMBL:Z98978; PIDN:CAB11682.1; GSPDB:GN00066;
A;Experimental source: strain 972h-; cosmid c27E2
A;Genetics: MUR1
R;Oliver, K.; Harris, D.; Wood, V.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21950
A;Accession: T40836
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-205,'P',207-1333 <OLI>
A;Cross-references: UNIPARC:UPI0000161F17; EMBL:AL023518; PIDN:CAB60245.1; GSPDB:GN00068
A;Experimental source: strain 972h-; cosmid c1020
A;Genetics: OLI1
R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21887
A;Accession: T39868
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-199,'I',201-695,'S',697-1333 <VOL>
A;Cross-references: UNIPARC:UPI000017B221; EMBL:AL021746; PIDN:CAA16837.1; GSPDB:GN00067
A;Experimental source: strain 972h-; cosmid c1E8
A;Genetics: VOL1
R;Connor, R.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1998
A;Reference number: Z21832
A;Accession: T50374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113,'P',115-205,'P',207-214,'F',216,'P',218-232,'I',234-273,'T',275-539,'I
A;Cross-references: UNIPARC:UPI000013529B; EMBL:AL021815; PIDN:CAB61532.1; GSPDB:GN00067
R;Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21733
A;Accession: T38069
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-50 <BRO>
A;Cross-references: UNIPARC:UPI0000069561; EMBL:Z98976; PIDN:CAB59613.1; GSPDB:GN00066;
A;Experimental source: strain 972h-; cosmid c1F2
C;Genetics: <MCL1>
A;Gene: SPAC26A3.13c; SPDB:SPAC1F2.03
A;Map position: 1
C;Genetics: <WED1>
A;Gene: SPAC9.04; SPAC9B6.02c
A;Map position: 2
C;Genetics: <VOL1>
A;Gene: retrotransposon
A;Map position: 2
A;Introns: 695/2
C;Genetics: <MUR1>
A;Gene: Tf2; SPAC27E2.08
A;Map position: 1
C;Genetics: <OLI1>
A;Gene: SPCC1020.14
A;Map position: 3

  Query Match 69.6%; Score 39; DB 2; Length 1333;
  Best Local Similarity 66.7%; Pred. No. 35;
  Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYPVGYFT 9
| |||||::
Db 731 KYYPVGYYS 739

RESULT 11
A90432
hypothetical protein SSO2588 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: A90432
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A90432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-142 <KUR>
A;Cross-references: UNIPROT:Q97VN1; UNIPARC:UPI00000647B3; GB:AE006641; NID:g13815891; P
C;Genetics:
A;Gene: SSO2588

  Query Match 66.1%; Score 37; DB 2; Length 142;
  Best Local Similarity 70.0%; Pred. No. 9;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIYPVGYFTK 10
| |||| ||:
Db 43 KDYPVGIFTE 52

RESULT 12
A81110
hypothetical protein NMB1202 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: A81110; G81906
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.,
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <TET>
A;Cross-references: UNIPROT:Q9JQW2; UNIPARC:UPI00000C4E54; GB:AE002468; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: G81906
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-153 <PAR>
A;Cross-references: UNIPARC:UPI00000C4E54; GB:AL162755; GB:AL157959; NID:g7379742; PIDN:
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMB1202; NMA1373
C;Superfamily: Neisseria meningitidis hypothetical protein NMB1202

  Query Match 66.1%; Score 37; DB 2; Length 153;
  Best Local Similarity 77.8%; Pred. No. 9.7;
  Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIYPVGYFT 9
|| | ||||
Db 40 KICPYGYFT 48

RESULT 13
G70482
pantothenate synthetase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: G70482
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: G70482
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-282 <AQF>
A;Cross-references: UNIPROT:O67891; UNIPARC:UPI00000567D7; GB:AE000774; NID:g2984324; P
A;Experimental source: strain VF5
C;Genetics:
A;Gene: panC
C;Superfamily: pantoate-beta-alanine ligase

  Query Match 66.1%; Score 37; DB 2; Length 282;
  Best Local Similarity 70.0%; Pred. No. 18;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIYPVGYFTK 10
:||| || ||
Db 98 EIYPKGYRTK 107

RESULT 14
D84280

hypothetical protein Vng1249c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84280
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84280
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <STO>
A;Cross-references: UNIPROT:Q9HQA7; UNIPARC:UPI000006383C; GB:AE004437; NID:g10580776; P
C;Genetics:
A;Gene: VNG1249C

Query Match 66.1%; Score 37; DB 2; Length 356;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 IYPVGYFT 9
              :||||:|:
Db        288 VYPVGFFS 295
```

RESULT 15
AF1593
B. subtilis PBSX prophage protein homolog lin1287 [imported] - Listeria innocua (strain
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF1593
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <GLA>
A;Cross-references: UNIPROT:Q92CA3; UNIPARC:UPI00000CC507; GB:AL592022; PIDN:CAC96518.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1287

Query Match 66.1%; Score 37; DB 2; Length 361;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KIYPVGY 7
              ||||:||
Db        209 KIYPLGY 215
```

Search completed: March 21, 2006, 21:49:28
Job time : 7.72727 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:58 ; Search time 29.6364 Seconds
(without alignments)
238.062 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	225	1	CDSN_PIG	O19084 sus scrofa
2	56	100.0	529	1	CDSN_HUMAN	Q15517 homo sapien
3	56	100.0	529	1	CDSN_PANTR	Q7yr44 pan troglod
4	56	100.0	534	1	CDSN_MACMU	Q5tm45 macaca mula
5	56	100.0	561	1	CDSN_MOUSE	Q7tpc1 mus musculu
6	46	82.1	249	1	PRRG_PRB01	Q9f7p4 gamma-prote
7	46	82.1	249	2	Q84C31_9BACT	Q84c31 uncultured
8	46	82.1	249	2	Q84C32_9BACT	Q84c32 uncultured
9	46	82.1	249	2	Q9AFF0_9BACT	Q9aff0 uncultured
10	46	82.1	249	2	Q9AFF2_9BACT	Q9aff2 uncultured
11	46	82.1	250	2	Q6J4G7_9BACT	Q6j4g7 uncultured
12	46	82.1	250	2	Q83TG9_9BACT	Q83tg9 uncultured
13	46	82.1	250	2	Q83TU0_9BACT	Q83tu0 uncultured
14	46	82.1	250	2	Q84C23_9BACT	Q84c23 uncultured
15	46	82.1	250	2	Q84C25_9BACT	Q84c25 uncultured
16	46	82.1	250	2	Q84C28_9BACT	Q84c28 uncultured
17	46	82.1	250	2	Q84C29_9BACT	Q84c29 uncultured
18	46	82.1	250	2	Q84C30_9BACT	Q84c30 uncultured
19	46	82.1	250	2	Q84BD7_9BACT	Q84bd7 uncultured
20	46	82.1	250	2	Q84BD8_9BACT	Q84bd8 uncultured
21	46	82.1	250	2	Q84BD9_9BACT	Q84bd9 uncultured
22	46	82.1	250	2	Q84BE0_9BACT	Q84be0 uncultured
23	46	82.1	250	2	Q84BE2_9BACT	Q84be2 uncultured
24	46	82.1	250	2	Q84BE3_9BACT	Q84be3 uncultured
25	46	82.1	250	2	Q84BE4_9BACT	Q84be4 uncultured
26	46	82.1	250	2	Q84BE5_9BACT	Q84be5 uncultured
27	46	82.1	250	2	Q84BE6_9BACT	Q84be6 uncultured
28	46	82.1	250	2	Q84BE7_9BACT	Q84be7 uncultured
29	46	82.1	250	2	Q84BE8_9BACT	Q84be8 uncultured
30	46	82.1	250	2	Q84BE9_9BACT	Q84be9 uncultured
31	46	82.1	251	2	Q6J5K4_9BACT	Q6j5k4 uncultured
32	46	82.1	251	2	Q99QB3_9BACT	Q99qb3 uncultured
33	46	82.1	251	2	Q9AFE5_9BACT	Q9afe5 uncultured
34	46	82.1	251	2	Q9AFE6_9BACT	Q9afe6 uncultured
35	46	82.1	251	2	Q9AFE7_9BACT	Q9afe7 uncultured
36	46	82.1	251	2	Q9AFE8_9BACT	Q9afe8 uncultured
37	46	82.1	251	2	Q9AFE9_9BACT	Q9afe9 uncultured
38	46	82.1	251	2	Q9AFF1_9BACT	Q9aff1 uncultured
39	46	82.1	251	2	Q9AFF3_9BACT	Q9aff3 uncultured
40	46	82.1	251	2	Q9AFF4_9BACT	Q9aff4 uncultured
41	46	82.1	251	2	Q9AFF5_9BACT	Q9aff5 uncultured
42	46	82.1	252	2	Q9AFG0_9BACT	Q9afg0 uncultured
43	45	80.4	252	2	Q9AFF9_9BACT	Q9aff9 uncultured
44	43	76.8	250	2	Q84C24_9BACT	Q84c24 uncultured
45	42	75.0	97	2	O12683_9VIRU	O12683 labroides d

ALIGNMENTS

RESULT 1
CDSN_PIG
ID CDSN_PIG STANDARD; PRT; 225 AA.
AC O19084;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Corneodesmosin (S protein) (Fragment).
GN Name=CDSN;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC TISSUE=Peripheral blood;
RA Velten F.W., Rogel-Gaillard C., Renard C., Pontarotti P.,
RA Tazi-Ahnini R., Vaiman M., Chardon P.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC intercellular structures that are involved in desquamation (By
CC similarity).
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---
DR EMBL; Z97402; CAB10690.1; -; Genomic_DNA.
FT COMPBIAS 79 85 Poly-Ser.
FT COMPBIAS 129 140 Poly-Ser.
FT COMPBIAS 159 168 Poly-Ser.
FT COMPBIAS 179 184 Poly-Ser.
FT NON_TER 1 1
FT NON_TER 225 225
SQ SEQUENCE 225 AA; 21644 MW; 8544070C41F7B280 CRC64;

Query Match 100.0%; Score 56; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIYPVGYFTK 10
||||||||||
Db 37 KIYPVGYFTK 46

RESULT 2
CDSN_HUMAN
ID CDSN_HUMAN STANDARD; PRT; 529 AA.
AC Q15517; O43509; Q5SQ85; Q5STD2; Q7LA70; Q7LA71; Q86Z04; Q8IZU4;
AC Q8IZU5; Q8IZU6; Q8N5P3; Q95IF9; Q9NP52; Q9NPE0; Q9NPG5; Q9NRH4;
AC Q9NRH5; Q9NRH6; Q9NRH7; Q9NRH8; Q9UBH8; Q9UIN6; Q9UIN7; Q9UIN8;

```
AC   Q9UIN9; Q9UIP0;
DT   15-JUL-1998 (Rel. 36, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LEU-18.
RC   TISSUE=Foreskin;
RX   MEDLINE=94022396; PubMed=8415725;
RA   Zhou Y., Chaplin D.D.;
RT   "Identification in the HLA class I region of a gene expressed late in
RT   keratinocyte differentiation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:9470-9474(1993).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT ALA-408.
RC   TISSUE=Epidermis;
RX   MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA   Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT   "Expression cloning of human corneodesmosin proves its identity with
RT   the product of the S gene and allows improved characterization of its
RT   processing during keratinocyte differentiation.";
RL   J. Biol. Chem. 273:22640-22647(1998).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS LEU-18; SER-153 DEL;
RP   PHE-202 AND LEU-410.
RX   MEDLINE=22256218; PubMed=12366786;
RX   DOI=10.1034/j.1399-0039.2002.600110.x;
RA   Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA   Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT   "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT   patients with psoriasis.";
RL   Tissue Antigens 60:77-83(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA   Shiina S., Tamiya G., Oka A., Inoko H.;
RT   "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA   Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RT   "Genome diversity in HLA: a new strategy for detection of genetic
RT   polymorphisms in expressed genes within the HLA class III and class I
RT   regions.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANTS LEU-18;
RP   ASN-143; SER-153 DEL; PHE-202; ALA-408; LEU-410 AND ASN-527.
RX   MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA   Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA   Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA   Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA   Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA   Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA   Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA   Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA   Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA   Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA   Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA   Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA   Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA   Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA   Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA   Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA   Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcott R.,
RA   Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA   Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA   Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA   Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA   Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA   Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLay K.,
RA   McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA   Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA   Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA   Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA   Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA   Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA   Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA   Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,
RA   Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA   Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA   Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT   "The DNA sequence and analysis of human chromosome 6.";
RL   Nature 425:805-811(2003).
RN   [7]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS LEU-18; SER-153
RP   DEL; PHE-202 AND LEU-410.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 30-529 (ALLELES 1.11; 1.21; 1.31;
RP   1.32; 1.41; 1.42; 1.43; 1.51; 1.52; 2.11; 2.21; 2.22 AND 2.23), AND
RP   VARIANTS PHE-56; ASN-143 DEL; ASN-150; PHE-202; GLY-401; ALA-408;
RP   LEU-410 AND ASN-527.
RX   MEDLINE=21108612; PubMed=11169256;
RX   DOI=10.1034/j.1399-0039.2001.057001032.x;
RA   Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT   "Identification of six novel polymorphisms in the human corneodesmosin
RT   gene.";
RL   Tissue Antigens 57:32-38(2001).
RN   [9]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 106-529, VARIANTS ASN-143;
RP   PHE-202; GLY-401; ALA-408 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX   MEDLINE=20066866; PubMed=10599883;
RX   DOI=10.1034/j.1399-0039.1999.540501.x;
RA   Jenisch S., Koch S., Henseler T., Nair R.P., Elder J.T., Watts C.E.,
RA   Westphal E., Voorhees J.J., Christophers E., Kroenke M.;
RT   "Corneodesmosin gene polymorphism demonstrates strong linkage
RT   disequilibrium with HLA and association with psoriasis vulgaris.";
RL   Tissue Antigens 54:439-449(1999).
RN   [10]
RP   PARTIAL PROTEIN SEQUENCE, AND CHARACTERIZATION.
RX   MEDLINE=98058975; PubMed=9395522; DOI=10.1074/jbc.272.50.31770;
RA   Simon M., Montezin M., Guerrin M., Durieux J.-J., Serre G.;
RT   "Characterization and purification of human corneodesmosin, an
RT   epidermal basic glycoprotein associated with corneocyte-specific
RT   modified desmosomes.";
RL   J. Biol. Chem. 272:31770-31776(1997).
RN   [11]
RP   INVOLVEMENT IN HTSS.
RX   PubMed=12754508; DOI=10.1038/ng1163;
```

```
RA   Levy-Nissenbaum E., Betz R.C., Frydman M., Simon M., Lahat H.,
RA   Bakhan T., Goldman B., Bygum A., Pierick M., Hillmer A.M., Jonca N.,
RA   Toribio J., Kruse R., Dewald G., Cichon S., Kubisch C., Guerrin M.,
RA   Serre G., Nothen M.M., Pras E.;
RT   "Hypotrichosis simplex of the scalp is associated with nonsense
RT   mutations in CDSN encoding corneodesmosin.";
RL   Nat. Genet. 34:151-153(2003).
RN   [12]
RP   VARIANTS PHE-202 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX   PubMed=10844560; DOI=10.1046/j.1523-1747.2000.00002.x;
RA   Enerbaeck C., Enlund F., Inerot A., Samuelsson L., Wahlstroem J.,
RA   Swanbeck G., Martinsson T.;
RT   "S gene (Corneodesmosin) diversity and its relationship to psoriasis;
RT   high content of cSNP in the HLA-linked S gene.";
RL   J. Invest. Dermatol. 114:1158-1163(2000).
RN   [13]
RP   VARIANTS PHE-202; GLY-401; ALA-408; LEU-410 AND ASN-527, AND
RP   ASSOCIATION WITH PSORIASIS.
RX   PubMed=12472658; DOI=10.1034/j.1399-0039.2002.600403.x;
RA   Orru S., Giuressi E., Casula M., Loizedda A., Murru R., Mulargia M.,
RA   Masala M.V., Cerimele D., Zucca M., Aste N., Biggio P., Carcassi C.,
RA   Contu L.;
RT   "Psoriasis is associated with a SNP haplotype of the corneodesmosin
RT   gene (CDSN).";
RL   Tissue Antigens 60:292-298(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation.
CC   -!- TISSUE SPECIFICITY: Exclusively expressed in skin.
CC   -!- POLYMORPHISM: Genetic variation in CDSN may be associated with
CC       susceptibility to psoriasis [MIM:177900]. Various CDSN alleles are
CC       known including alleles 1.11, 1.21, 1.31, 1.32, 1.41, 1.42, 1.43,
CC       1.51, 1.52, 2.11, 2.21, 2.22 and 2.23.
CC   -!- DISEASE: Defects in CDSN are a cause of hypotrichosis simplex of
CC       the scalp (HTSS) [MIM:146520]; also known as hypotrichosis Spanish
CC       type. HTSS is an autosomal dominant form of isolated alopecia.
CC       Affected individuals have normal hair in early childhood but
CC       experience progressive loss of scalp hair beginning in the middle
CC       of the first decade and almost complete baldness by the third
CC       decade.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 501.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L20815; AAA21321.1; ALT_SEQ; mRNA.
DR   EMBL; AF030130; AAC24196.1; -; mRNA.

  Query Match             100.0%;  Score 56;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 3
CDSN_PANTR
ID   CDSN_PANTR     STANDARD;      PRT;   529 AA.
AC   Q7YR44;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BA000041; BAC78170.1; -; Genomic_DNA.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    529       Corneodesmosin.
FT   COMPBIAS     62    464       Ser-rich.
FT   CARBOHYD    172    172       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 4
CDSN_MACMU
ID   CDSN_MACMU     STANDARD;      PRT;   534 AA.
AC   Q5TM45;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=CDSN;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15269276; DOI=10.1093/molbev/msh216;
RA   Kulski J.K., Anzai T., Shiina T., Inoko H.;
RT   "Rhesus macaque class I duplicon structures, organization, and
RT   evolution within the alpha block of the major histocompatibility
RT   complex.";
RL   Mol. Biol. Evol. 21:2079-2091(2004).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB128049; BAD69749.1; -; Genomic_DNA.
KW   Signal.
```

```
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    534       Corneodesmosin.
FT   COMPBIAS     62    469       Ser-rich.
SQ   SEQUENCE   534 AA;  51983 MW;  CB6A9DF256C32874 CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        322 KIYPVGYFTK 331


RESULT 5
CDSN_MOUSE
ID   CDSN_MOUSE     STANDARD;      PRT;   561 AA.
AC   Q7TPC1;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=Cdsn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C3H/He; TISSUE=Osteoblast;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BC055373; AAH55373.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039518; Mus musculus.
DR   MGI; MGI:3505689; Cdsn.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    561       Corneodesmosin.
FT   COMPBIAS     62    493       Ser-rich.
SQ   SEQUENCE   561 AA;  54414 MW;  7A6EBEF364038746 CRC64;

  Query Match             100.0%;  Score 56;  DB 1;  Length 561;
  Best Local Similarity   100.0%;  Pred. No. 0.039;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        344 KIYPVGYFTK 353


RESULT 6
PRRG_PRB01
ID   PRRG_PRB01     STANDARD;      PRT;   249 AA.
AC   Q9F7P4;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Green-light absorbing proteorhodopsin precursor (GPR).
OS   Gamma-proteobacterium EBAC31A08.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; environmental samples.
OX   NCBI_TaxID=133804;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20446260; PubMed=10988064; DOI=10.1126/science.289.5486.1902;
RA   Beja O., Aravind L., Koonin E.V., Suzuki M.T., Hadd A., Nguyen L.P.,
RA   Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,
RA   DeLong E.F.;
RT   "Bacterial rhodopsin: evidence for a new type of phototrophy in the
RT   sea.";
RL   Science 289:1902-1906(2000).
RN   [2]
RP   CHARACTERIZATION OF PROTON RELEASE CYCLE, AND MUTAGENESIS OF CYS-107;
RP   CYS-156 AND CYS-175.
RX   PubMed=11943070; DOI=10.1186/1472-6793-2-5;
RA   Krebs R.A., Alexiev U., Partha R., DeVita A., Braiman M.S.;
RT   "Detection of fast light-activated H+ release and M intermediate
RT   formation from proteorhodopsin.";
RL   BMC Physiol. 2:5-5(2002).
RN   [3]
RP   CHARACTERIZATION OF PHOTOCHEMICAL CYCLE, AND MUTAGENESIS OF ASP-97 AND
RP   GLU-108.
RX   MEDLINE=21965984; PubMed=11969395; DOI=10.1021/bi025563x;
RA   Dioumaev A.K., Brown L.S., Shih J., Spudich E.N., Spudich J.L.,
RA   Lanyi J.K.;
RT   "Proton transfers in the photochemical reaction cycle of
RT   proteorhodopsin.";
RL   Biochemistry 41:5348-5358(2002).
RN   [4]
RP   CHARACTERIZATION OF PHOTOCHEMICAL CYCLE.
RX   MEDLINE=22434728; PubMed=12547799;
RA   Varo G., Brown L.S., Lakatos M., Lanyi J.K.;
RT   "Characterization of the photochemical reaction cycle of
RT   proteorhodopsin.";
RL   Biophys. J. 84:1202-1207(2003).
RN   [5]
RP   COMPARISON WITH PHOTOCHEMICAL CYCLE OF BLUE PROTEORHODOPSIN.
RX   PubMed=12821661; DOI=10.1074/jbc.M305716200;
RA   Wang W.W., Sineshchekov O.A., Spudich E.N., Spudich J.L.;
RT   "Spectroscopic and photochemical characterization of a deep ocean
RT   proteorhodopsin.";
RL   J. Biol. Chem. 278:33985-33991(2003).
RN   [6]
RP   PRELIMINARY CHARACTERIZATION OF PUMP VECTORIALITY.
RX   MEDLINE=22195742; PubMed=12206764; DOI=10.1016/S0022-2836(02)00696-4;
RA   Friedrich T., Geibel S., Kalmbach R., Chizhov I., Ataka K.,
RA   Heberle J., Engelhard M., Bamberg E.;
RT   "Proteorhodopsin is a light-driven proton pump with variable
RT   vectoriality.";
RL   J. Mol. Biol. 321:821-838(2002).
RN   [7]
RP   CHARACTERIZATION OF PUMP VECTORIALITY.
RX   MEDLINE=22604715; PubMed=12719254;
RA   Lakatos M., Lanyi J.K., Szakacs J., Varo G.;
RT   "The photochemical reaction cycle of proteorhodopsin at low pH.";
RL   Biophys. J. 84:3252-3256(2003).
RN   [8]
```

```
RP   CHARACTERIZATION OF PUMP VECTORIALITY.
RX   MEDLINE=22652026; PubMed=12767242; DOI=10.1021/bi034253r;
RA   Dioumaev A.K., Wang J.M., Balint Z., Varo G., Lanyi J.K.;
RT   "Proton transport by proteorhodopsin requires that the retinal Schiff
RT   base counterion Asp-97 be anionic.";
RL   Biochemistry 42:6582-6587(2003).
RN   [9]
RP   ENVIRONMENTAL DISTRIBUTION.
RX   MEDLINE=21351675; PubMed=11459054; DOI=10.1038/35081051;
RA   Beja O., Spudich E.N., Spudich J.L., Leclerc M., DeLong E.F.;
RT   "Proteorhodopsin phototrophy in the ocean.";
RL   Nature 411:786-789(2001).
RN   [10]
RP   ENVIRONMENTAL DISTRIBUTION, AND MUTAGENESIS OF VAL-68 AND LEU-105.
RX   MEDLINE=22568278; PubMed=12682005; DOI=10.1093/emboj/cdg183;
RA   Man D., Wang W., Sabehi G., Aravind L., Post A.F., Massana R.,
RA   Spudich E.N., Spudich J.L., Beja O.;
RT   "Diversification and spectral tuning in marine proteorhodopsins.";
RL   EMBO J. 22:1725-1731(2003).
CC   -!- FUNCTION: Light-driven proton pump that generates photothrophic
CC       energy.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- MISCELLANEOUS: It presents a fast proton release and an
CC       alkaliphilic photocycle, consistent with its marine origin and the
CC       near-surface environment where this bacterium was collected.
CC       Transport occurs only at pHs above 7 and is unidirectional.
CC   -!- MISCELLANEOUS: Mutagenesis experiments that removed all the Cys
CC       residues of GPR show that at least one of the Cys residues in the
CC       protein is probably modified post-translationally.
CC   -!- SIMILARITY: Belongs to the archaeal/bacterial/fungal opsin family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AF279106; AAG10475.1; -; Genomic_DNA.
DR   HSSP; P02945; 1C3W.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
DR   PROSITE; PS00327; BACTERIAL_OPSIN_RET; FALSE_NEG.
KW   Chromophore; Hydrogen ion transport; Ion transport;
KW   Photoreceptor protein; Receptor; Retinal protein;
KW   Sensory transduction; Signal; Transmembrane; Transport.
FT   SIGNAL        1     17       Potential.
FT   CHAIN        18    249       Green-light absorbing proteorhodopsin.
FT   TRANSMEM     30     49       Potential.
FT   TRANSMEM     62     84       Potential.
FT   TRANSMEM     99    121       Potential.
FT   TRANSMEM    128    147       Potential.
FT   TRANSMEM    151    168       Potential.
FT   TRANSMEM    189    211       Potential.
FT   TRANSMEM    221    243       Potential.
FT   BINDING     231    231       Retinal chromophore (covalent) (By
FT                                similarity).
FT   SITE         97     97       Primary proton acceptor.
FT   SITE        105    105       Responsible for spectral tuning.
FT   SITE        108    108       Primary proton donor.
FT   MUTAGEN      68     68       V->I: No effect.
FT   MUTAGEN      97     97       D->E: Slight changes in spectral shifts;
FT                                more rapid formation of the M
FT                                intermediate.
FT   MUTAGEN      97     97       D->N: Elimination of the pH-induced shift
FT                                and stabilization of the absorption
FT                                spectrum in the acidic form.
FT   MUTAGEN     105    105       L->Q: Absorbs blue light; slower
FT                                photocycle.
FT   MUTAGEN     107    107       C->S: Lower molecular weight on SDS gels
FT                                potentially due to failure to bind either
FT                                sugar or lipid; when associated with S-
FT                                156 and S-175.
FT   MUTAGEN     108    108       E->Q: Changes in the photocycle.
FT   MUTAGEN     156    156       C->S: Lower molecular weight on SDS gels
FT                                potentially due to failure to bind either
FT                                sugar or lipid; when associated with S-
FT                                107 and S-175.
FT   MUTAGEN     175    175       C->S: Lower molecular weight on SDS gels
FT                                potentially due to failure to bind either
FT                                sugar or lipid; when associated with S-
FT                                107 and S-156.
SQ   SEQUENCE   249 AA;  27251 MW;  5D6B54FB36C237EA CRC64;

  Query Match             82.1%;  Score 46;  DB 1;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


RESULT 7
Q84C31_9BACT
ID   Q84C31_9BACT PRELIMINARY;      PRT;   249 AA.
AC   Q84C31;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22873787; PubMed=14510837;
RX   DOI=10.1046/j.1462-2920.2003.00493.x;
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RT   "Novel Proteorhodopsin variants from the Mediterranean and Red Seas.";
RL   Environ. Microbiol. 5:842-849(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15122363;
RA   Man-Aharonovich D., Sabehi G., Sineshchekov O.A., Spudich E.N.,
RA   Spudich J.L., Beja O.;
RT   "Characterization of RS29, a blue-green proteorhodopsin variant from
RT   the Red Sea.";
RL   Photochem. Photobiol. Sci. 3:459-462(2004).
DR   EMBL; AY250715; AAO73909.1; -; Genomic_DNA.
DR   EMBL; AY598756; AAT09901.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER       1      1
FT   NON_TER     249    249
SQ   SEQUENCE   249 AA;  27334 MW;  AE08F287BE25A818 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 8
Q84C32_9BACT
ID   Q84C32_9BACT PRELIMINARY;      PRT;   249 AA.
```

```
AC   Q84C32;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22873787; PubMed=14510837;
RX   DOI=10.1046/j.1462-2920.2003.00493.x;
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RT   "Novel Proteorhodopsin variants from the Mediterranean and Red Seas.";
RL   Environ. Microbiol. 5:842-849(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY250714; AAO73908.2; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER     249    249
SQ   SEQUENCE   249 AA;  27348 MW;  A978F287BE25A818 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 9
Q9AFF0_9BACT
ID   Q9AFF0_9BACT PRELIMINARY;      PRT;   249 AA.
AC   Q9AFF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21351675; PubMed=11459054; DOI=10.1038/35081051;
RA   Beja O., Spudich E.N., Spudich J.L., Leclerc M., DeLong E.F.;
RT   "Proteorhodopsin phototrophy in the ocean.";
RL   Nature 411:786-789(2001).
DR   EMBL; AF349989; AAK30187.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER       1      1
FT   NON_TER     249    249
SQ   SEQUENCE   249 AA;  27182 MW;  2766D196BFEBA167 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        201 IYPVGYFT 208


RESULT 10
Q9AFF2_9BACT
ID   Q9AFF2_9BACT PRELIMINARY;      PRT;   249 AA.
AC   Q9AFF2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21351675; PubMed=11459054; DOI=10.1038/35081051;
RA   Beja O., Spudich E.N., Spudich J.L., Leclerc M., DeLong E.F.;
RT   "Proteorhodopsin phototrophy in the ocean.";
RL   Nature 411:786-789(2001).
DR   EMBL; AF349987; AAK30185.1; -; Genomic_DNA.
DR   HSSP; P02945; 1M0L.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER       1      1
FT   NON_TER     249    249
SQ   SEQUENCE   249 AA;  27194 MW;  59FE3E78FE5A2717 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        201 IYPVGYFT 208


RESULT 11
Q6J4G7_9BACT
ID   Q6J4G7_9BACT PRELIMINARY;      PRT;   250 AA.
AC   Q6J4G7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22873787; PubMed=14510837;
RX   DOI=10.1046/j.1462-2920.2003.00493.x;
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RT   "Novel Proteorhodopsin variants from the Mediterranean and Red Seas.";
RL   Environ. Microbiol. 5:842-849(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15466697; DOI=10.1073/pnas.0403999101;
RA   Bielawski J.P., Dunn K.A., Sabehi G., Beja O.;
RT   "Darwinian adaptation of proteorhodopsin to different light
RT   intensities in the marine environment.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14824-14829(2004).
DR   EMBL; AY601905; AAT12291.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
```

```
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27308 MW;  5FDB6DFB3209C7EA CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 12
Q83TG9_9BACT
ID   Q83TG9_9BACT PRELIMINARY;      PRT;   250 AA.
AC   Q83TG9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22568278; PubMed=12682005; DOI=10.1093/emboj/cdg183;
RA   Man D., Wang W., Sabehi G., Aravind L., Post A.F., Massana R.,
RA   Spudich E.N., Spudich J.L., Beja O.;
RT   "Diversification and spectral tuning in marine proteorhodopsins.";
RL   EMBO J. 22:1725-1731(2003).
DR   EMBL; AY210900; AAO21436.1; -; Genomic_DNA.
DR   EMBL; AY210906; AAO21442.1; -; Genomic_DNA.
DR   HSSP; P02945; 1M0L.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27204 MW;  DD5B1381F9E7D2A4 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 13
Q83TU0_9BACT
ID   Q83TU0_9BACT PRELIMINARY;      PRT;   250 AA.
AC   Q83TU0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22568278; PubMed=12682005; DOI=10.1093/emboj/cdg183;
RA   Man D., Wang W., Sabehi G., Aravind L., Post A.F., Massana R.,
RA   Spudich E.N., Spudich J.L., Beja O.;
RT   "Diversification and spectral tuning in marine proteorhodopsins.";
RL   EMBO J. 22:1725-1731(2003).
DR   EMBL; AY210907; AAO21443.1; -; Genomic_DNA.
DR   EMBL; AY210911; AAO21447.1; -; Genomic_DNA.
DR   EMBL; AY210912; AAO21448.1; -; Genomic_DNA.
DR   EMBL; AY210903; AAO21439.1; -; Genomic_DNA.
DR   EMBL; AY210901; AAO21437.1; -; Genomic_DNA.
DR   HSSP; P02945; 1M0L.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27260 MW;  AFCB7DFB3209DB87 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 14
Q84C23_9BACT
ID   Q84C23_9BACT PRELIMINARY;      PRT;   250 AA.
AC   Q84C23;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22873787; PubMed=14510837;
RX   DOI=10.1046/j.1462-2920.2003.00493.x;
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RT   "Novel Proteorhodopsin variants from the Mediterranean and Red Seas.";
RL   Environ. Microbiol. 5:842-849(2003).
DR   EMBL; AY250723; AAO73917.1; -; Genomic_DNA.
DR   HSSP; P02945; 1C8R.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   InterPro; IPR000238; Rib_bind_factA.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
DR   PROSITE; PS01319; RBFA; UNKNOWN_1.
FT   NON_TER       1      1
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27326 MW;  9B06454178ED951F CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 15
Q84C25_9BACT
ID   Q84C25_9BACT PRELIMINARY;      PRT;   250 AA.
```

```
AC   Q84C25;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Proteorhodopsin (Fragment).
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22873787; PubMed=14510837;
RX   DOI=10.1046/j.1462-2920.2003.00493.x;
RA   Sabehi G., Massana R., Bielawski J.P., Rosenberg M., Delong E.F.,
RA   Beja O.;
RT   "Novel Proteorhodopsin variants from the Mediterranean and Red Seas.";
RL   Environ. Microbiol. 5:842-849(2003).
DR   EMBL; AY250721; AAO73915.1; -; Genomic_DNA.
DR   HSSP; P02945; 1C8R.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005216; F:ion channel activity; IEA.
DR   GO; GO:0006811; P:ion transport; IEA.
DR   InterPro; IPR001425; Bac_rhodopsin.
DR   Pfam; PF01036; Bac_rhodopsin; 1.
DR   PRINTS; PR00251; BACTRLOPSIN.
DR   PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
FT   NON_TER       1      1
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27291 MW;  73735EEECC70D205 CRC64;

  Query Match             82.1%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


Search completed: March 21, 2006, 21:48:44
Job time : 29.6364 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:40:11 ; Search time 8124 Seconds
(without alignments)
11125.195 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1539	96.8	1590	6	AR429544	AR429544 Sequence
	2	1488	93.6	1725	6	AX786044	AX786044 Sequence
	3	1488	93.6	1725	8	AF030130	AF030130 Homo sapi
	4	1434	90.2	1590	6	A84909	A84909 Sequence 2
	5	1434	90.2	1590	6	BD072818	BD072818 Polypepti
	6	1355	85.2	1764	8	AF224751	AF224751 Homo sapi
	7	1355	85.2	1764	8	AF224753	AF224753 Homo sapi
	8	1355	85.2	2127	8	AF224752	AF224752 Homo sapi
	9	1355	85.2	4740	6	AX786043	AX786043 Sequence
	10	1355	85.2	5337	6	BD095296	BD095296 The metho
	11	1355	85.2	5700	6	AX233490	AX233490 Sequence
	12	1355	85.2	25235	6	BD095297	BD095297 The metho
c	13	1355	85.2	30911	8	AB088114	AB088114 Homo sapi
	14	1355	85.2	39196	8	AB023060	AB023060 Homo sapi
	15	1355	85.2	44379	8	AC006163	AC006163 Homo sapi
	16	1355	85.2	55050	6	AX522117	AX522117 Sequence
c	17	1355	85.2	56809	8	CR759805	CR759805 Human DNA
	18	1355	85.2	110000	8	BA000025_08	Continuation (9 of
c	19	1355	85.2	135396	8	BX927139	BX927139 Human DNA
c	20	1355	85.2	349980	6	CS039418	CS039418 Sequence
	21	1353	85.1	1506	6	AX786098	AX786098 Sequence
	22	1350	84.9	1509	8	AF491329S2	AF491330 Homo sapi
	23	1335	84.0	1725	6	CQ722442	CQ722442 Sequence
	24	1304	82.0	1764	8	AF224747	AF224747 Homo sapi
	25	1304	82.0	2127	8	AF286165	AF286165 Homo sapi
c	26	1304	82.0	81418	8	CR753819	CR753819 Human DNA
	27	1253	78.8	1764	8	AF224748	AF224748 Homo sapi
	28	1253	78.8	1764	8	AF224749	AF224749 Homo sapi
	29	1253	78.8	1764	8	AF224754	AF224754 Homo sapi
c	30	1253	78.8	113388	8	AL773544	AL773544 Human DNA
	31	1245	78.3	2530	6	AX336639	AX336639 Sequence
	32	1245	78.3	2530	6	AX786041	AX786041 Sequence
	33	1245	78.3	2530	8	HUMSPROT	L20815 Human S pro
	34	1123	70.6	1324	8	HSA238461	AJ238461 Homo sapi
	35	1072	67.4	1324	8	HSA238462	AJ238462 Homo sapi
	36	1072	67.4	1324	8	HSA238463	AJ238463 Homo sapi
	37	1021	64.2	1324	8	HSA238464	AJ238464 Homo sapi
	38	1021	64.2	1324	8	HSA238465	AJ238465 Homo sapi
	39	1010	63.5	1761	8	AF224750	AF224750 Homo sapi
	40	978	61.5	2284	6	CS086650	CS086650 Sequence
	41	896	56.4	110000	8	BA000041_03	Continuation (4 of
	42	894	56.2	1767	8	AF224757	AF224757 Homo sapi
	43	894	56.2	1767	8	AF224758	AF224758 Homo sapi
	44	843	53.0	1767	8	AF224756	AF224756 Homo sapi
c	45	843	53.0	204298	8	AL662844	AL662844 Human DNA
c	46	792	49.8	3889	8	HSM807103	BX640954 Homo sapi
	47	733	46.1	1587	11	BT019904	BT019904 Synthetic
	48	704	44.3	1764	8	AF224755	AF224755 Homo sapi
c	49	704	44.3	70288	8	AL662867	AL662867 Human DNA
	50	702	44.2	1327	8	HSA238466	AJ238466 Homo sapi
	51	699	44.0	1516	8	AF491325S2	AF491326 Homo sapi
	52	684	43.0	2062	8	BC031993	BC031993 Homo sapi
	53	648	40.8	1516	8	AF491327S2	AF491328 Homo sapi
	54	632	39.7	1324	8	HSA238467	AJ238467 Homo sapi
	55	621	39.1	1590	6	AX233491	AX233491 Sequence
	56	322	20.3	783	10	BV661298	BV661298 S217P6643
	57	300	18.9	402	8	AY328475	AY328475 Homo sapi
	58	255	16.0	404	10	BV184125	BV184125 sqnm14318
	59	249	15.7	402	8	AY328474	AY328474 Homo sapi
	60	247	15.5	580	10	BV188296	BV188296 sqnm15795
	61	237	14.9	516	10	BV188297	BV188297 sqnm15795
	62	150	9.4	201	10	BV205968	BV205968 sqnm21963
	63	150	9.4	201	10	BV205975	BV205975 sqnm21965
	64	139	8.7	294	8	HSL14A30	U18169 Human HLA c
	65	119	7.5	110000	8	AB128049_13	Continuation (14 o
c	66	119	7.5	200068	8	AC148691	AC148691 Macaca mu
	67	105	6.6	201	10	BV199621	BV199621 sqnm20178
c	68	103	6.5	175689	8	AC148703	AC148703 Macaca mu
	69	87	5.5	87	8	AF491327S1	AF491327 Homo sapi
	70	87	5.5	87	8	AF491329S1	AF491329 Homo sapi
c	71	86	5.4	6810	8	CR936880	CR936880 Human DNA
	72	85	5.3	99	6	AX786097	AX786097 Sequence
	73	80	5.0	1007	10	G07214	G07214 human STS W
	74	79	5.0	79	6	AX233523	AX233523 Sequence
	75	51	3.2	87	8	AF491325S1	AF491325 Homo sapi

ALIGNMENTS

RESULT 1
AR429544
LOCUS AR429544 1590 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6645509.
ACCESSION AR429544
VERSION AR429544.1 GI:40189837
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1590)

Applicant

```
  AUTHORS   Serre,G.B.R., Simon,M. and Weber-Vivat,M.
  TITLE     Polypeptide expressed in the horny layer of epidermis and use
            thereof
  JOURNAL   Patent: US 6645509-A 2 11-NOV-2003;
            Societe L'Oreal S.A.; Paris;
            FRX;
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             96.8%;  Score 1539;  DB 6;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1589;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 2
AX786044
LOCUS       AX786044               1725 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequence 14 from Patent WO03050311.
ACCESSION   AX786044
VERSION     AX786044.1  GI:32953661
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
  TITLE     Methods and compositions for diagnosing and treating hypotrichosis
            simplex
  JOURNAL   Patent: WO 03050311-A 14 19-JUN-2003;
            Wella AG (DE)
FEATURES             Location/Qualifiers
     source          1. .1725
                     /organism="Homo sapiens"
```

```
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             93.6%;  Score 1488;  DB 6;  Length 1725;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1588;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

```
RESULT 3
AF030130
LOCUS       AF030130               1725 bp    mRNA    linear   PRI 03-AUG-2000
DEFINITION  Homo sapiens corneodesmosin mRNA, complete cds.
ACCESSION   AF030130
VERSION     AF030130.1  GI:2921269
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1725)
  AUTHORS   Guerrin,M., Simon,M., Montezin,M., Haftek,M., Vincent,C. and
            Serre,G.
  TITLE     Expression cloning of human corneodesmosin proves its identity with
            the product of the S gene and allows improved characterization of
            its processing during keratinocyte differentiation
  JOURNAL   J. Biol. Chem. 273 (35), 22640-22647 (1998)
   PUBMED   9712893
REFERENCE   2  (bases 1 to 1725)
  AUTHORS   Guerrin,M., Simon,M., Montezin,M., Vincent,C. and Serre,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-OCT-1997) LBPDM, Faculte de Medecine Purpan, 37
            Allees Jules Guesde, Toulouse 31000, France
FEATURES             Location/Qualifiers
     source          1. .1725
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21"
                     /cell_type="keratinocyte; corneocyte"
                     /tissue_type="epidermis"
     CDS             15. .1604
                     /note="associated with corneocyte-specific modified
                     desmosomes; late differentiation protein of epidermis;
                     similar to human differentiated keratinocyte S protein
                     encoded by GenBank Accession Number L20815"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAC24196.1"
                     /db_xref="GI:2921270"
                     /translation="MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSS
                     FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPC
                     SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
                     GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSP
                     GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
                     PSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGS
                     KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
                     LGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             93.6%;  Score 1488;  DB 8;  Length 1725;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1588;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604


RESULT 4
A84909
LOCUS       A84909                 1590 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 2 from Patent WO9844105.
ACCESSION   A84909
VERSION     A84909.1  GI:6733756
KEYWORDS    .
```

```
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 1590)
  AUTHORS   Simon,M. and Serre,G.B.
  TITLE     POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS AND USE
            THEREOF
  JOURNAL   Patent: WO 9844105-A 2 08-OCT-1998;
            OREAL (FR); SIMON MICHEL (FR)
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             90.2%;  Score 1434;  DB 6;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1534;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174

Qy        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234

Qy        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294

Qy        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        355 TCCGGATCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474

Qy        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534

Qy        535 AATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        535 AATGACAACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594

Qy        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654

Qy        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714

Qy        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774

Qy        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834

Qy        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894

Qy        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954

Qy        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014

Qy       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074

Qy       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134

Qy       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194

Qy       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254

Qy       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314

Qy       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374

Qy       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434

Qy       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554

Qy       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||||||||
Db       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 5
BD072818
LOCUS       BD072818                1590 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Polypeptide expressed in the horny layer of epidermis and use
            thereof.
ACCESSION   BD072818
VERSION     BD072818.1  GI:22618421
KEYWORDS    JP 2001511019-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1590)
  AUTHORS   Serre,G.B.R., Simon,M. and Vivat,M.W.
  TITLE     Polypeptide expressed in the horny layer of epidermis and use
  JOURNAL   Patent: JP 2001511019-A 1 07-AUG-2001;
            L'OREAL
COMMENT     OS   Unidentified
            PN   JP 2001511019-A/1
            PD   07-AUG-2001
            PF   27-MAR-1998 JP 1998541244
            PR   28-MAR-1997 FR     97/03899,11-SEP-1997 FR     97/11317 PI
```

```
            GUY BRUNO RENE SERRE,MICHEL SIMON,MARINA WEBER VIVAT PC
            C12N15/12,C07K14/47,A61K38/17,C07K16/18
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Polypeptide expressed in the horny layer of epidermis and use
            CC            thereof
            FH   Key             Location/Qualifiers
            FT   source          1. .1590
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             90.2%;  Score 1434;  DB 6;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1534;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174

Qy        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234

Qy        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294

Qy        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        355 TCCGGATCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474

Qy        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534

Qy        535 AATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        535 AATGACAACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594

Qy        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654

Qy        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714

Qy        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774

Qy        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834

Qy        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894

Qy        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954

Qy        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014

Qy       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074

Qy       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134

Qy       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194

Qy       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254

Qy       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314

Qy       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374

Qy       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434

Qy       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554

Qy       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||||||||
Db       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 6
AF224751
LOCUS       AF224751               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.41 allele, partial
            cds.
ACCESSION   AF224751
VERSION     AF224751.1  GI:8699031
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
```

```
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.41"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.41"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.41"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78563.1"
                     /db_xref="GI:8699032"
                     /translation="TLAKSIGTPSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match           85.2%;  Score 1355;  DB 8;  Length 1764;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1603 GTCCATAA 1610

RESULT 7
```

```
AF224753
LOCUS       AF224753                1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.51 allele, partial
            cds.
ACCESSION   AF224753
VERSION     AF224753.1  GI:8699035
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.51"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.51"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.51"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78565.1"
                     /db_xref="GI:8699036"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 8;  Length 1764;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
             ||||||||
Db       1603 GTCCATAA 1610


RESULT 8
AF224752
LOCUS       AF224752               2127 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.42 allele, partial
            cds.
ACCESSION   AF224752
VERSION     AF224752.2  GI:9027570
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2127)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 2127)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
REFERENCE   3  (bases 1 to 2127)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
  REMARK    Sequence update by submitter
COMMENT     On Jul 11, 2000 this sequence version replaced gi:8699033.
FEATURES             Location/Qualifiers
     source          1. .2127
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <471. .>2127
                     /gene="CDSN"
                     /allele="CDSN-1.42"
     mRNA            <471. .>2127
                     /gene="CDSN"
                     /allele="CDSN-1.42"
                     /product="corneodesmosin"
     CDS             <471. .1973
                     /gene="CDSN"
                     /allele="CDSN-1.42"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78564.1"
                     /db_xref="GI:8699034"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match           85.2%;  Score 1355;  DB 8;  Length 2127;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 525

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 585

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 645

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 705

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        706 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 765

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 825

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 885

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        886 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 945

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 1005

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 1065

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 1125

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 1185

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 1245

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 1305

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1365
```

```
Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1425

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1485

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1545

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1605

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1665

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1725

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1785

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1845

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1905

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1965

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1966 GTCCATAA 1973
```

```
RESULT 9
AX786043
LOCUS       AX786043               4740 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequence 13 from Patent WO03050311.
ACCESSION   AX786043
VERSION     AX786043.1  GI:32953660
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
  TITLE     Methods and compositions for diagnosing and treating hypotrichosis
            simplex
  JOURNAL   Patent: WO 03050311-A 13 19-JUN-2003;
            Wella AG (DE)
FEATURES             Location/Qualifiers
     source          1. .4740
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 6;  Length 4740;
  Best Local Similarity   99.8%;  Pred. No. 0;
```

```
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3039 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3098

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3099 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3158

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3159 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3218

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3219 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3278

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3279 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3338

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3339 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3398

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3399 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3458

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3459 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3518

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3519 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3578

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3579 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3638

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3639 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3698

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3699 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3758

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3759 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3818

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3819 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3878

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3879 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3938

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3939 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3998

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3999 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4058

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4059 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4118
```

```
Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4119 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4178

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4179 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4238

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4239 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4298

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4299 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4358

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4359 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4418

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4419 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4478

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4479 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4538

Qy       1583 GTCCATAA 1590
             ||||||||
Db       4539 GTCCATAA 4546
```

RESULT 10
BD095296

LOCUS	BD095296	5337 bp DNA linear PAT 27-AUG-2002
DEFINITION	The method of testing for psoriasis vulgaris.	
ACCESSION	BD095296	
VERSION	BD095296.1 GI:22640884	
KEYWORDS	WO 0142458-A/1.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
REFERENCE	1 (bases 1 to 5337)	
AUTHORS	Inoko,H. and Tamiya,G.	
TITLE	The method of testing for psoriasis vulgaris	
JOURNAL	Patent: WO 0142458-A 1 14-JUN-2001;	
	HIDETOSHI INOKO,GEN TAMIYA	
COMMENT	OS Homo sapiens (human)	
	PN WO 0142458-A/1	
	PD 14-JUN-2001	
	PF 06-DEC-2000 WO 2000JP008624	
	PR 06-DEC-1999 JP 99P 346867	
	PI HIDETOSHI INOKO,GEN TAMIYA	
	PC C12N15/12,C12Q1/68	
	CC The method of testing for psoriasis vulgaris	
	FH Key Location/Qualifiers	
	FT exon (1). .(99)	
	FT exon (2904). .(5337).	
FEATURES	Location/Qualifiers	
source	1. .5337	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		

Query Match 85.2%; Score 1355; DB 6; Length 5337;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
                ||||||||
Db       4401 GTCCATAA 4408


RESULT 11
AX233490
LOCUS       AX233490                5700 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 133 from Patent WO0162788.
ACCESSION   AX233490
VERSION     AX233490.1  GI:15592975
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Olaveson,M., Lench,N., Allen,M. and Tazi-Ahnini,R.U.
  TITLE     Corneodesmosin based test and model for inflammatory disease
  JOURNAL   Patent: WO 0162788-A 133 30-AUG-2001;
            Oxagen Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .5700
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     5'UTR           251. .498
     exon            393. .497
     3'UTR           393. .497
     exon            3301. .4807
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 6;  Length 5700;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3299 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3358

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3359 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3418

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3419 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3478

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3479 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3538

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3539 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3598

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3599 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3658

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3659 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3718

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3719 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3778

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3779 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3838

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3839 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3898

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3899 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3958

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3959 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 4018

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4019 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 4078

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4079 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 4138

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4139 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 4198

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 4258

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4318

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4378

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4379 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4438

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4439 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4498

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       4499 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4558

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4559 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4618

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4619 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4678

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4679 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4738

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4739 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4798

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4799 GTCCATAA 4806


RESULT 12
BD095297
LOCUS       BD095297               25235 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  The method of testing for psoriasis vulgaris.
ACCESSION   BD095297
VERSION     BD095297.1  GI:22640885
KEYWORDS    WO 0142458-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 25235)
  AUTHORS   Inoko,H. and Tamiya,G.
  TITLE     The method of testing for psoriasis vulgaris
  JOURNAL   Patent: WO 0142458-A 2 14-JUN-2001;
            HIDETOSHI INOKO,GEN TAMIYA
COMMENT     OS   Homo sapiens (human)
            PN   WO 0142458-A/2
            PD   14-JUN-2001
            PF   06-DEC-2000 WO 2000JP008624
            PR   06-DEC-1999 JP 99P   346867
            PI   HIDETOSHI INOKO,GEN TAMIYA
            PC   C12N15/12,C12Q1/68
            CC   The method of testing for psoriasis vulgaris
            FH   Key             Location/Qualifiers
            FT   exon            (1). .(420)
            FT   exon            (1282). .(1405)
            FT   exon            (1602). .(1702)
            FT   exon            (1602). .(1631)
            FT   exon            (2352). .(2364)
            FT   exon            (6287). .(6509)
            FT   exon            (10417). .(10493)
            FT   exon            (14244). .(14407)
            FT   exon            (14244). .(14344)
            FT   exon            (25190). .(25235).
FEATURES            Location/Qualifiers
     source         1. .25235
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match            85.2%;  Score 1355;  DB 6;  Length 25235;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691
```

```
Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
             ||||||||
Db      24052 GTCCATAA 24059


RESULT 13
AB088114/c
LOCUS       AB088114               30911 bp    DNA     linear   PRI 08-JAN-2003
DEFINITION  Homo sapiens STG, S, SPR1, SEEK1 genes.
ACCESSION   AB088114
VERSION     AB088114.1  GI:27544424
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shiina,T., Ota,M., Katsuyama,Y., Hashimoto,N. and Inoko,H.
  TITLE     Genome diversity in HLA: A new strategy for detection of genetic
            polymorphisms in expressed genes within the HLA class III and class
            I regions
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 30911)
  AUTHORS   Shiina,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUL-2002) Takashi Shiina, Tokai University School of
            Medicine, Molecular Life Sciense 2; Bohseidai, Isehara, Kanagawa
            259-1193, Japan (E-mail:tshiina@is.icc.u-tokai.ac.jp,
            Tel:81-463-93-1121, Fax:81-463-94-8884)
FEATURES             Location/Qualifiers
     source          1. .30911
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
                     /cell_line="LKT3"
                     /cell_type="B cell"
     exon            1317. .2383
                     /number=2
     gene            complement(join(1473. .2383,2581. .2647))
                     /gene="STG"
     CDS             complement(join(1473. .2383,2581. .2647))
                     /gene="STG"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC54947.1"
                     /db_xref="GI:27544425"
                     /translation="MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNLGTNL
                     PQLGQPSSTGPSNSEHPQPALDPRSNDLARVPLKLSVPASDGFPPAGGSAVQRWPPSW
                     GLPAMDSWPPEDPWQMMAAAAEDRLGEALPEELSYLSSAAALAPGSGPLPGESSPDAT
                     GLSPKASLLHQDSESRRLPRSNSLGAGGKILSQRPPWSLIHRVLPDHPWGTLNPSVSW
                     GGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGSWGNINRYPGGSWGNIN
                     RYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG"
     exon            2581. .2647
                     /number=1
     exon            4938. .4983
                     /number=1
     exon            5185. .7618
                     /number=2
     gene            complement(join(6114. .7618,10423. .10459))
                     /gene="S"
     CDS             complement(join(6114. .7618,10423. .10459))
                     /gene="S"
                     /codon_start=1
                     /product="modified desmosin expressed in non-cornified
                     epidermis"
                     /protein_id="BAC54948.1"
                     /db_xref="GI:27544426"
                     /translation="MMALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG
                     DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYS
                     SGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSAL
                     PTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVS
                     HSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKS
                     YGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKY
                     FSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSS
                     GSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLT
                     LTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGE
                     LLDSP"
     exon            10423. .10521
                     /number=1
     exon            15765. .15928
                     /number=2
     exon            19676. .19752
                     /number=3
     exon            27641. .28413
                     /number=2
     gene            complement(join(28058. .28413,29079. .29133))
                     /gene="SPR1"
     CDS             complement(join(28058. .28413,29079. .29133))
                     /gene="SPR1"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC54949.1"
                     /db_xref="GI:27544427"
                     /translation="MILNWKLLGILVLCLHTRGISGSEGHPSHPPAEDREEAGSPTLP
                     QGPPVPGDPWPGAPPLFEDPPPTRPSRPWRDLPETGVWPPEPPRTDPPQPPRPDDPWP
                     AGPQPPENPWPPAPEVDNRPQEEPDLDPPREEYR"
     exon            28537. .28566
                     /number=4
     exon            28763. .28886
                     /number=5
     gene            join(28864. .28886,29748. .30039)
                     /gene="SEEK1"
     CDS             join(28864. .28886,29748. .30039)
                     /gene="SEEK1"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAC54950.1"
                     /db_xref="GI:27544428"
                     /translation="MEPANHFWHAGDLQAMISKEFHLAATQDDCRKGRTQEDILVPSS
                     HPELFASVLPMAPEEAARLQQPQPLPPPSGIHLSASRTLAPTLLYSSPPSHSPFGLSS
                     LI"
     exon            29079. .29448
                     /gene="SEEK1"
                     /number=1
     exon            29748. .30167
                     /number=6
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 8;  Length 30911;
```

```
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7621 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 7562

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7561 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 7502

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7501 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 7442

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7441 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 7382

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       7381 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 7322

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7321 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 7262

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 7202

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       7201 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 7142

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7141 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 7082

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7081 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 7022

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7021 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 6962

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6961 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 6902

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6901 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 6842

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6841 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 6782

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6781 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 6722

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6721 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 6662

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6661 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 6602

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       6601 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 6542

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6541 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 6482

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6481 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 6422

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6421 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 6362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6361 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 6302

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 6242

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6241 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 6182

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6181 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 6122

Qy       1583 GTCCATAA 1590
              ||||||||
Db       6121 GTCCATAA 6114
```

```
RESULT 14
AB023060
LOCUS       AB023060               39196 bp    DNA     linear   PRI 20-NOV-1999
DEFINITION  Homo sapiens genomic DNA, chromosome 6p21.3, HLA class I region,
            clone:TY1C2, complete sequence.
ACCESSION   AB023060
VERSION     AB023060.1  GI:5672629
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Shiina,T., Tamiya,G., Oka,A., Takishima,N., Yamagata,T.,
            Kikkawa,E., Iwata,K., Tomizawa,M., Okuaki,N., Kuwano,Y.,
            Watanabe,K., Fukuzumi,Y., Itakura,S., Sugawara,C., Ono,A.,
            Yamazaki,M., Tashiro,H., Ando,A., Ikemura,T., Soeda,E., Kimura,M.,
            Bahram,S. and Inoko,H.
  TITLE     Molecular dynamics of MHC genesis unraveled by sequence analysis of
            the 1,796,938-bp HLA class I region
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (23), 13282-13287 (1999)
   PUBMED   10557312
REFERENCE   2  (bases 1 to 39196)
  AUTHORS   Shiina,T. and Takishima,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-1999) Takashi Shiina, Tokai University School of
            Medicine, Department of Molecular Life Science 2; Bohseidai,
            Isehara, Kanagawa 259-1193, Japan
            (E-mail:tshiina@is.icc.u-tokai.ac.jp, Tel:81-463-93-1121,
            Fax:81-463-94-8884)
FEATURES             Location/Qualifiers
     source          1. .39196
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
```

```
                    /clone="TY1C2"
                    /cell_line="BOLETH"
                    /tissue_type="B cell line"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 8;  Length 39196;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31433 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 31492

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31493 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 31552

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31553 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 31612

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31613 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 31672

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31673 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 31732

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31733 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 31792

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31793 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 31852

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      31853 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 31912

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31913 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 31972

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31973 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 32032

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32033 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 32092

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32093 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 32152

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32153 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 32212

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32213 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 32272

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32273 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 32332

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32333 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 32392

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32393 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 32452

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32453 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 32512

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32513 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 32572

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32573 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 32632

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32633 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 32692

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32693 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 32752

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32753 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 32812

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32813 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 32872

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32873 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 32932

Qy       1583 GTCCATAA 1590
              ||||||||
Db      32933 GTCCATAA 32940


RESULT 15
AC006163
LOCUS       AC006163               44379 bp    DNA     linear   PRI 08-DEC-1998
DEFINITION  Homo sapiens clone UWGC:y24c203 from 6p21, complete sequence.
ACCESSION   AC006163
VERSION     AC006163.1  GI:3980468
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 44379)
  AUTHORS   Janer,M., Guillaudeux,T., Vu,Q., Kutyavin,T., Harter,H. and
            Geraghty,D.E.
  TITLE     Large scale sequence analysis of the human MHC class I region
  JOURNAL   Unpublished (1998)
  REMARK    Fred Hutchinson Cancer Research Center
            The Clinical Research Division
            1100  Fairview Ave. N., P.O. Box 19024
            Seattle, WA 98109-1024
REFERENCE   2  (bases 1 to 44379)
  AUTHORS   Geraghty,D.E. and Olson,M.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1998) Human Genome Center, University of
            Washington, Box 352145, Seattle, WA 98195, USA
  REMARK    University of Washington Human Genome Center
              Box 352145, Seattle, WA 98195
              Contact: Daniel E. Geraghty (geraghty@fhcrc.org)
COMMENT     Overlapping Sequences:
               5': UWGC:y24c027
               3': UWGC:y28c284
```

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 75.7%
DS or two chemistry coverage: 99.8%
Single stranded regions: 3

Sequence Validation:
This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments below a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

BglII		HindIII		NsiI	
Map	Seq	Map	Seq	Map	Seq
1387.80	1395.00	12090.89	11940.00	4644.97	4536.00
4231.92	4186.00	812.42	792.00	5445.21	5321.00
6249.71	6247.00	3308.27	3248.00	8990.18	8852.00
6238.00	6287.00	16520.91	16559.00	7230.63	7121.00
				1823.37	1834.00
				1393.29	1387.00
				1130.32	1115.00
				1491.44	1517.00
				7227.23	7286.00
				1320.82	1318.00

```
FEATURES             Location/Qualifiers
     source          1. .44379
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21"
                     /clone="CGM1:A194D6"
                     /sub_clone="UWGC:y24c203"
                     /cell_line="CGM1"
                     /clone_lib="Wash U YAC library"
     repeat_region   complement(2553. .2834)
                     /rpt_family="Alu"
     repeat_region   complement(4746. .5027)
                     /rpt_family="Alu"
     repeat_region   complement(5607. .5741)
                     /rpt_family="MER3"
     repeat_region   complement(7610. .7914)
                     /rpt_family="Alu"
     repeat_region   complement(7975. .8066)
                     /rpt_family="MIR"
     repeat_region   8408. .8688
                     /rpt_family="Alu"
     repeat_region   complement(9622. .9703)
                     /rpt_family="Alu"
     repeat_region   13938. .14427
                     /rpt_family="MLT1"
     repeat_region   18018. .18301
                     /rpt_family="Alu"
     repeat_region   18781. .19668
                     /rpt_family="Alu"
     repeat_region   19952. .20029
                     /rpt_family="MIR"
     repeat_region   complement(20556. .20664)
                     /rpt_family="Alu"
     repeat_region   complement(20837. .21131)
                     /rpt_family="Alu"
     repeat_region   complement(21665. .21952)
                     /rpt_family="Alu"
     repeat_region   22020. .22306
                     /rpt_family="Alu"
     repeat_region   complement(22367. .22965)
                     /rpt_family="Alu"
     repeat_region   23430. .23514
                     /rpt_family="Alu"
     repeat_region   complement(24023. .24085)
                     /rpt_family="MER43"
     repeat_region   complement(24360. .24638)
                     /rpt_family="Alu"
     repeat_region   complement(24926. .25161)
                     /rpt_family="MER43"
     repeat_region   25402. .25700
                     /rpt_family="Alu"
     repeat_region   complement(25883. .26147)
                     /rpt_family="Alu"
     repeat_region   26244. .26492
                     /rpt_family="MER25"
     repeat_region   26974. .27230
                     /rpt_family="L1"
     repeat_region   complement(27383. .27671)
                     /rpt_family="Alu"
     repeat_region   27907. .28188
                     /rpt_family="Alu"
     repeat_region   29056. .29348
                     /rpt_family="Alu"
     repeat_region   complement(29958. .30257)
                     /rpt_family="Alu"
     repeat_region   complement(31194. .31464)
                     /rpt_family="Alu"
     repeat_region   complement(31726. .31993)
                     /rpt_family="Alu"
     repeat_region   33393. .33658
                     /rpt_family="Alu"
     repeat_region   34321. .34463
                     /rpt_family="MIR"
     repeat_region   complement(34482. .35051)
                     /rpt_family="Alu"
     repeat_region   complement(35136. .35177)
                     /rpt_family="MIR"
     repeat_region   complement(35574. .35688)
                     /rpt_family="Alu"
     repeat_region   complement(36649. .36906)
                     /rpt_family="Alu"
     repeat_region   37317. .37611
                     /rpt_family="Alu"
     repeat_region   complement(37673. .37921)
                     /rpt_family="Alu"
     repeat_region   complement(38353. .38622)
                     /rpt_family="Alu"
     repeat_region   complement(38822. .38950)
                     /rpt_family="Alu"
     repeat_region   39166. .39433
                     /rpt_family="Alu"
     repeat_region   complement(39578. .39725)
                     /rpt_family="Alu"
     repeat_region   complement(41088. .41375)
                     /rpt_family="Alu"
     repeat_region   42842. .43123
                     /rpt_family="Alu"
```

```
     repeat_region   complement(44135. .44372)
                     /rpt_family="Alu"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 8;  Length 44379;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9999 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 10058

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10059 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 10118

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10119 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 10178

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10179 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 10238

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10239 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 10298

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10299 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 10358

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10359 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 10418

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      10419 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 10478

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10479 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 10538

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10539 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 10598

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10599 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 10658

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10659 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 10718

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10719 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 10778

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10779 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 10838

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10839 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 10898

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10899 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 10958

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10959 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 11018

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11019 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 11078

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11079 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 11138

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11139 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 11198

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11199 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 11258

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11259 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 11318

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11319 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 11378

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11379 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 11438

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11439 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 11498

Qy       1583 GTCCATAA 1590
              ||||||||
Db      11499 GTCCATAA 11506


RESULT 16
AX522117
LOCUS       AX522117               55050 bp    DNA     linear   PAT 24-OCT-2002
DEFINITION  Sequence 4 from Patent WO02064831.
ACCESSION   AX522117
VERSION     AX522117.1  GI:24411002
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lench,N.J., Allen,M.J. and Nicholls,R.K.
  TITLE     Test and model for inflammatory disease
  JOURNAL   Patent: WO 02064831-A 4 22-AUG-2002;
            Oxagen Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .55050
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 6;  Length 55050;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47898 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 47957

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
```

```
Db      47958 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 48017

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48018 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 48077

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48078 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 48137

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      48138 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 48197

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48198 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 48257

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48258 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 48317

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      48318 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 48377

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48378 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 48437

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48438 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 48497

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48498 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 48557

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48558 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 48617

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48618 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 48677

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48678 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 48737

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48738 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 48797

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48798 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 48857

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48858 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 48917

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48918 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 48977

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48978 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 49037

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      49038 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 49097

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49098 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 49157

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49158 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 49217

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49218 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 49277

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49278 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 49337

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49338 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 49397

Qy       1583 GTCCATAA 1590
              ||||||||
Db      49398 GTCCATAA 49405
```

RESULT 17
CR759805/c

```
LOCUS       CR759805              56809 bp    DNA     linear   PRI 02-APR-2005
DEFINITION  Human DNA sequence from clone DAMC-11N4 on chromosome 6, complete
            sequence.
ACCESSION   CR759805
VERSION     CR759805.5  GI:62176980
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 56809)
  AUTHORS   Mashreghi-Mohammadi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 2, 2005 this sequence version replaced gi:55468475.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
```

```
          the MHC Haplotype Consortium and collaborators. Further information
          can be found at
          http://www.sanger.ac.uk/HGP/Chr6/MHC
          DAMC-11N4 is from the DNA-Arts.org BAC library MCF.1 VECTOR:
          pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .56809
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAMC-11N4"
                     /clone_lib="DNA-Arts.org BAC library MCF.1"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 8;  Length 56809;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52434 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 52375

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52374 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 52315

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52314 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 52255

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52254 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 52195

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      52194 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 52135

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52134 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 52075

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52074 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 52015

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      52014 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 51955

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51954 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 51895

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51894 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 51835

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51834 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 51775

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51774 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 51715

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51714 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 51655

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      51654 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 51595

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51594 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 51535

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51534 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 51475

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51474 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 51415

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51414 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 51355

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51354 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 51295

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51294 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 51235

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51234 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 51175

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51174 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 51115

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51114 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 51055

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51054 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 50995

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      50994 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 50935

Qy       1583 GTCCATAA 1590
              ||||||||
Db      50934 GTCCATAA 50927
```

```
RESULT 18
BA000025_08
WPCOMMENT
  Sequence split into 23 fragments  LOCUS BA000025 Accession BA000025
     Fragment Name      Begin     End
     BA000025_00            1    110000
     BA000025_01       100001    210000
     BA000025_02       200001    310000
     BA000025_03       300001    410000
     BA000025_04       400001    510000
     BA000025_05       500001    610000
     BA000025_06       600001    710000
     BA000025_07       700001    810000
     BA000025_08       800001    910000
     BA000025_09       900001   1010000
     BA000025_10      1000001   1110000
     BA000025_11      1100001   1210000
     BA000025_12      1200001   1310000
     BA000025_13      1300001   1410000
     BA000025_14      1400001   1510000
     BA000025_15      1500001   1610000
     BA000025_16      1600001   1710000
```

```
     BA000025_17       1700001   1810000
     BA000025_18       1800001   1910000
     BA000025_19       1900001   2010000
     BA000025_20       2000001   2110000
     BA000025_21       2100001   2210000
     BA000025_22       2200001   2229817
Continuation (9 of 23) of BA000025 from base 800001 (BA000025 Homo sapiens genomic DNA,

  Query Match             85.2%;  Score 1355;  DB 8;  Length 110000;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24567 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 24626

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24627 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 24686

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24687 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 24746

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24747 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 24806

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      24807 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 24866

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24867 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 24926

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24927 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 24986

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      24987 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 25046

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25047 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 25106

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25107 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 25166

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25167 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 25226

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25227 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 25286

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25287 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 25346

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25347 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 25406

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25407 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 25466

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25467 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 25526

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25527 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 25586

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25587 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 25646

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25647 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 25706

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25707 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 25766

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25767 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 25826

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25827 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 25886

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25887 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 25946

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25947 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 26006

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26007 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 26066

Qy       1583 GTCCATAA 1590
              ||||||||
Db      26067 GTCCATAA 26074


RESULT 19
BX927139/c
LOCUS       BX927139              135396 bp    DNA     linear   PRI 20-OCT-2004
DEFINITION  Human DNA sequence from clone DAMA-72F24 on chromosome 6, complete
            sequence.
ACCESSION   BX927139
VERSION     BX927139.8  GI:54311982
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 135396)
  AUTHORS   Phillimore,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 20, 2004 this sequence version replaced gi:48374126.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
```

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6/MHC
DAMA-72F24 is from the DNA-Arts human BAC library MANN.1 VECTOR: pBeloBAC11.

FEATURES Location/Qualifiers
source 1. .135396
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAMA-72F24"
/clone_lib="DNA-Arts BAC library MANN.1"

ORIGIN

Query Match 85.2%; Score 1355; DB 8; Length 135396;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109053 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 108994

Qy       143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108993 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 108934

Qy       203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108933 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 108874

Qy       263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108873 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 108814

Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    108813 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 108754

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108753 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 108694

Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108693 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 108634

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    108633 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 108574

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108573 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 108514

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108513 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 108454

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108453 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 108394

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108393 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 108334

Qy       803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108333 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 108274

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108273 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 108214

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108213 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 108154

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108153 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 108094

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108093 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 108034

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108033 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 107974

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107973 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 107914

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107913 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 107854

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107853 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 107794

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107793 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 107734

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107733 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 107674

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107673 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 107614

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107613 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 107554

Qy      1583 GTCCATAA 1590
             ||||||||
Db    107553 GTCCATAA 107546
```

RESULT 20
CS039418/c
LOCUS CS039418 349980 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 228 from Patent WO2005019477.
ACCESSION CS039418

```
VERSION     CS039418.1  GI:61847431
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lewin,J., Berlin,K., Hildmann,T., Olek,A., Beck,S. and Novik,K.
  TITLE     Methods and compositions for differentiating tissues or cell types
            using epigenetic markers
  JOURNAL   Patent: WO 2005019477-A 228 03-MAR-2005;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="unknown base
                     seq 205 to long, 4.267.840
                     replaced by
                     new seq 205, from 0.000.001 to 0.349.980
                     new seq 222, from 0.300.001 to 0.649.980
                     new seq 223, from 0.600.001 to 0.949.980
                     new seq 224, from 0.900.001 to 1.249.980
                     new seq 225, from 1.200.001 to 1.549.980
                     new seq 226, from 1.500.001 to 1.849.980
                     new seq 227, from 1.800.001 to 2.149.980
                     new seq 228, from 2.100.001 to 2.449.980
                     new seq 229, from 2.400.001 to 2.749.980
                     new seq 230, from 2.700.001 to 3.049.980
                     new seq 231, from 3.000.001 to 3.349.980
                     new seq 232, from 3.300.001 to 3.649.980
                     new seq 233, from 3.600.001 to 3.949.980
                     new seq 234, from 3.900.001 to 4.249.980
                     new seq 235, from 4.200.001 to 4.549.980
                     new seq 236, from 4.500.001 to 4.647.455"
ORIGIN

  Query Match             85.2%;  Score 1355;  DB 6;  Length 349980;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278914 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 278855

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278854 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 278795

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278794 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 278735

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278734 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 278675

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
               ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      278674 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 278615

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278614 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 278555

Qy         443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278554 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 278495

Qy         503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
               ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      278494 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 278435

Qy         563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278434 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 278375

Qy         623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278374 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 278315

Qy         683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278314 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 278255

Qy         743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278254 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 278195

Qy         803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278194 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 278135

Qy         863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278134 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 278075

Qy         923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278074 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 278015

Qy         983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      278014 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 277955

Qy        1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277954 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 277895

Qy        1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277894 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 277835

Qy        1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277834 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 277775

Qy        1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277774 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 277715

Qy        1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277714 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 277655

Qy        1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277654 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 277595

Qy        1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277594 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 277535

Qy        1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277534 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 277475

Qy        1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277474 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 277415

Qy        1583 GTCCATAA 1590
               ||||||||
Db      277414 GTCCATAA 277407
```

```
RESULT 21
AX786098
LOCUS       AX786098               1506 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequence 68 from Patent WO03050311.
ACCESSION   AX786098
VERSION     AX786098.1  GI:32953663
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
  TITLE     Methods and compositions for diagnosing and treating hypotrichosis
            simplex
  JOURNAL   Patent: WO 03050311-A 68 19-JUN-2003;
            Wella AG (DE)
FEATURES             Location/Qualifiers
     source          1. .1506
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             85.1%;  Score 1353;  DB 6;  Length 1506;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          85 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 60

Qy         145 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 120

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 180

Qy         265 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 240

Qy         325 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCA 384
               |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         241 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCA 300

Qy         385 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 444
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 360

Qy         445 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 420

Qy         505 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAAC 564
               |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         421 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAAC 480

Qy         565 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 624
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 540

Qy         625 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 684
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 600

Qy         685 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 660

Qy         745 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 804
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 720

Qy         805 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 864
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 780

Qy         865 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 924
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 840

Qy         925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 900

Qy         985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 960

Qy        1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1020

Qy        1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1080

Qy        1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        1081 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCTCC 1140

Qy        1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1200

Qy        1285 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1260

Qy        1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1320

Qy        1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1380

Qy        1465 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1440

Qy        1525 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1500

Qy        1585 CCATAA 1590
               ||||||
Db        1501 CCATAA 1506


RESULT 22
AF491329S2
LOCUS       AF491329S2             1509 bp    DNA     linear   PRI 15-NOV-2002
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.44t allele, exon 2
            and complete cds.
ACCESSION   AF491330
VERSION     AF491330.1  GI:25005090
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1509)
  AUTHORS   Hui,J., Oka,A., Tamiya,G., Tomizawa,M., Kulski,J.K., Penhale,W.J.,
            Tay,G.K., Iizuka,M., Ozawa,A. and Inoko,H.
  TITLE     Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
            patients with psoriasis
  JOURNAL   Tissue Antigens 60 (1), 77-83 (2002)
   PUBMED   12366786
REFERENCE   2  (bases 1 to 1509)
  AUTHORS   Inoko,H., Tamiya,G., Oka,A., Hui,J., Kulski,J. and Tay,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Pathology, UWA, QEII Medical Centre,
            Nedlands, WA 6009, Australia
FEATURES             Location/Qualifiers
     source          1. .1509
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            order(AF491329.1:<1. .87,1. .>1503)
                     /gene="CDSN"
                     /allele="1.44t"
     mRNA            join(AF491329.1:<1. .87,1. .>1503)
                     /gene="CDSN"
                     /allele="1.44t"
                     /product="corneodesmosin"
     CDS             join(AF491329.1:1. .87,1. .1503)
                     /gene="CDSN"
                     /allele="1.44t"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAN70996.1"
                     /db_xref="GI:25005091"
                     /translation="MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSS
                     FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPC
                     SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
                     GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYPTKENPVKGSP
                     GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
                     PSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGS
                     KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
                     LGPQLADPEVFLPQGELLDSP"
     exon            1. .>1503
                     /gene="CDSN"
                     /allele="1.44t"
                     /number=2
ORIGIN

  Query Match             84.9%;  Score 1350;  DB 8;  Length 1509;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1500;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 60

Qy        148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 120

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 180

Qy        268 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 240

Qy        328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCATCC 300

Qy        388 GGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGA 360

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAA 420

Qy        508 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACCCT 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAACCCT 480

Qy        568 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCCAA 540

Qy        628 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGA 600

Qy        688 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGT 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGT 660

Qy        748 CAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAA 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAA 720

Qy        808 GGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTA 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTA 780

Qy        868 GACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCA 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCA 840

Qy        928 GGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCT 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCT 900

Qy        988 GTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAA 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAA 960

Qy       1048 TACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTC 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTC 1020

Qy       1108 CAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGA 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGA 1080

Qy       1168 CCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGT 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCTCCGGT 1140

Qy       1228 TCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGC 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 TCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGC 1200

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGC 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGC 1260

Qy       1348 AGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGG 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGG 1320

Qy       1408 GGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1467
```

```
Db       1321 GGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGC 1380

Qy       1468 AGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTG 1527

Db       1381 AGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTG 1440

Qy       1528 GGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCA 1587

Db       1441 GGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCA 1500

Qy       1588 TAA 1590

Db       1501 TAA 1503


RESULT 23
CQ722442
LOCUS       CQ722442               1725 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8376 from Patent WO02068579.
ACCESSION   CQ722442
VERSION     CQ722442.1  GI:42283299
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 8376 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1725
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           84.0%;  Score 1335;  DB 6;  Length 1725;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1585;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Db        615 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Db       1215 GGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

```
Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604


RESULT 24
AF224747
LOCUS       AF224747               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.11 allele, partial
            cds.
ACCESSION   AF224747
VERSION     AF224747.1  GI:8699023
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.11"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.11"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.11"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78559.1"
                     /db_xref="GI:8699024"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             82.0%;  Score 1304;  DB 8;  Length 1764;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1504;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1242
```

```
Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
             ||||||||
Db       1603 GTCCATAA 1610
```

```
RESULT 25
AF286165
LOCUS       AF286165               2127 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.43 allele, partial
            cds.
ACCESSION   AF286165
VERSION     AF286165.1  GI:9937326
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2127)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 2127)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) LBPDM, Faculte de Medecine, 37, allees J.
            Guesde, Toulouse cedex 31073, France
FEATURES             Location/Qualifiers
     source          1. .2127
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <1. .>2127
                     /gene="CDSN"
                     /allele="1.43"
     intron          <1. .468
                     /gene="CDSN"
                     /allele="1.43"
                     /number=1
     mRNA            <469. .>2127
                     /gene="CDSN"
                     /allele="1.43"
                     /product="corneodesmosin"
     exon            469. .>2127
                     /gene="CDSN"
                     /allele="1.43"
                     /number=2
     CDS             <469. .1973
                     /gene="CDSN"
                     /allele="1.43"
                     /codon_start=3
                     /product="corneodesmosin"
                     /protein_id="AAG02419.1"
                     /db_xref="GI:9937327"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
     3'UTR           1974. .>2127
                     /gene="CDSN"
                     /allele="1.43"
ORIGIN

  Query Match             82.0%;  Score 1304;  DB 8;  Length 2127;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1504;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 525

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 585

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 645

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 705

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        706 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 765

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 825

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 885

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        886 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 945

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 1005

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 1065

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 1125

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
```

```
Db       1126 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 1185

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862

Db       1186 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 1245

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922

Db       1246 CTGTAGACAAATCCTATGGTGGCTATGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 1305

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982

Db       1306 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1365

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042

Db       1366 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1425

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102

Db       1426 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1485

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162

Db       1486 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1545

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222

Db       1546 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1605

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282

Db       1606 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1665

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342

Db       1666 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1725

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402

Db       1726 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1785

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462

Db       1786 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1845

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522

Db       1846 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1905

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582

Db       1906 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1965

Qy       1583 GTCCATAA 1590

Db       1966 GTCCATAA 1973


RESULT 26
CR753819/c
LOCUS       CR753819               81418 bp    DNA     linear   PRI 06-OCT-2004
DEFINITION  Human DNA sequence from clone DADB-14104 on chromosome 6, complete
            sequence.
ACCESSION   CR753819
VERSION     CR753819.5  GI:53850278
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 81418)
  AUTHORS   Mashreghi-Mohammadi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 6, 2004 this sequence version replaced gi:52839457.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            DADB-14104 is from a DNA-arts DBB human bac library VECTOR:
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .81418
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DADB-14104"
                     /clone_lib="DNA-arts-BAC.1-DBB.1"
ORIGIN

  Query Match             82.0%;  Score 1304;  DB 8;  Length 81418;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1504;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142

Db       38797 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 38738

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202

Db       38737 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 38678

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262

Db       38677 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 38618

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322

Db       38617 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 38558

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382

Db       38557 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 38498

Qy         383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
```

```
Db      38497 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 38438

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502

Db      38437 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 38378

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562

Db      38377 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 38318

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622

Db      38317 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 38258

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682

Db      38257 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 38198

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742

Db      38197 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 38138

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802

Db      38137 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 38078

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862

Db      38077 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 38018

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922

Db      38017 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 37958

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982

Db      37957 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 37898

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042

Db      37897 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 37838

Qy       1043 GCAAATACTTCTCCAGCAACCCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102

Db      37837 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 37778

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162

Db      37777 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 37718

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222

Db      37717 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 37658

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282

Db      37657 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 37598

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342

Db      37597 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 37538

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402

Db      37537 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 37478

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462

Db      37477 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 37418

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522

Db      37417 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 37358

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582

Db      37357 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 37298

Qy       1583 GTCCATAA 1590

Db      37297 GTCCATAA 37290
```

```
RESULT 27
AF224748
LOCUS       AF224748               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.21 allele, partial
            cds.
ACCESSION   AF224748
VERSION     AF224748.1  GI:8699025
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.21"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.21"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.21"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78560.1"
                     /db_xref="GI:8699026"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSGSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             78.8%;  Score 1253;  DB 8;  Length 1764;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
```

```
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202

Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262

Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322

Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382

Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442

Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502

Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562

Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622

Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682

Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742

Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802

Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862

Db        823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922

Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982

Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042

Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102

Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162

Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222

Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCGGTTCTAGCATTTCCAGCAGCT 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282

Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342

Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402

Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462

Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522

Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582

Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590

Db       1603 GTCCATAA 1610
```

```
RESULT 28
AF224749
LOCUS       AF224749               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.31 allele, partial
            cds.
ACCESSION   AF224749
VERSION     AF224749.1  GI:8699027
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.31"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.31"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.31"
```

```
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78561.1"
                     /db_xref="GI:8699028"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCFTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match           78.8%;  Score 1253;  DB 8;  Length 1764;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCTTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1603 GTCCATAA 1610


RESULT 29
AF224754
LOCUS       AF224754               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.52 allele, partial
            cds.
ACCESSION   AF224754
VERSION     AF224754.1  GI:8699037
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.52"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-1.52"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-1.52"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78566.1"
                     /db_xref="GI:8699038"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             78.8%;  Score 1253;  DB 8;  Length 1764;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||| ||||||||||| ||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGACCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 822

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 882

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 942

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        943 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCACCAAAGAGA 1002

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1062

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1122

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1182

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1242

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1302

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1362

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1422

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1482

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1542

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1602

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1603 GTCCATAA 1610


RESULT 30
AL773544/c
```

```
LOCUS       AL773544              113388 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone DAQB-185I1 on chromosome 6 Contains
            the C6orf15 gene for chromosome 6 open reading frame 15, the CDSN
            gene for corneodesmosin, the PSORS1C1 gene for psoriasis
            susceptibility 1 candidate 1, the PSORS1C2 gene for psoriasis
            susceptibility 1 candidate 2, the POLR2LP pseudogene for polymerase
            (RNA) II (DNA directed) polypeptide L, the C6orf18 gene for
            chromosome 6 open reading frame 18, the TCF19 gene for
            transcription factor 19 (SC1), the POU5F1 gene for POU domain,
            class 5, transcription factor 1, the PSORS1C3 gene for psoriasis
            susceptibility 1 candidate 3, and 1 CpG island, complete sequence.
ACCESSION   AL773544
VERSION     AL773544.5  GI:37805611
KEYWORDS    HTG; C6orf15; C6orf18; CDSN; CpG island; POLR2LP; POU5F1; PSORS1C1;
            PSORS1C2; PSORS1C3; TCF19.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 113388)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 21, 2003 this sequence version replaced gi:22759515.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            DAQB-185I1 is from a DNA-arts QBL human bac library VECTOR:
            pBeloBAC11
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .113388
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAQB-185I1"
                     /clone_lib="DNA-arts-BAC.1-QBL.1"
     misc_feature    1
                     /note="Clone_left_end: DAQB-185I1"
     misc_feature    7991
                     /note="Clone_right_end: DAQB-218D23"
     gene            complement(33031. .34363)
                     /gene="C6orf15"
                     /locus_tag="DAQB-185I1.6-001"
     mRNA            complement(join(33031. .34099,34297. .34363))
                     /gene="C6orf15"
                     /locus_tag="DAQB-185I1.6-001"
                     /product="chromosome 6 open reading frame 15"
                     /note="match: cDNAs: AB031481.1 Em:AY358438.1"
     polyA_site      complement(33031)
                     /gene="C6orf15"
                     /locus_tag="DAQB-185I1.6-001"
     polyA_signal    complement(33044. .33049)
                     /gene="C6orf15"
                     /locus_tag="DAQB-185I1.6-001"
     CDS             complement(join(33189. .34099,34297. .34363))
                     /gene="C6orf15"
                     /locus_tag="DAQB-185I1.6-001"
                     /standard_name="OTTHUMP00000014814"
                     /note="match: proteins: Q9BGL9 Q9UIG3"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 15"
                     /protein_id="CAI18478.1"
                     /db_xref="GI:56206476"
                     /db_xref="UniProt/TrEMBL:Q5SQ81"
                     /translation="MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTDL
                     PQLGQPSSTGPSNSEHPQPALDPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSW
                     GLPAMDSWPPEDPWQMMAAAAEDRLGEALPEELSYLSSAAALAPGSGPLPGESSPDAT
                     GLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRPPWSLIHRVLPDHPWGTLNPSVSW
                     GGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGSWGNINRYPGGSWGNIN
                     RYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG"
     gene            join(36639. .36699,47489. .47652,51402. .51478,60266. .60295,
                     60492. .60614,61476. .61894)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-001"
     mRNA            join(36639. .36699,47489. .47652,51402. .51478,60266. .60295,
                     60492. .60614,61476. .61894)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-001"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: cDNAs: AB031479.1 Em:AF484418.1"
     gene            join(36639. .36699,61476. .61872)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-003"
     mRNA            join(36639. .36699,61476. .61872)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-003"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: Em:AI343394.1
                     match: cDNAs: Em:AF484419.1"
     gene            complement(join(36896. .39334,42146. .42257))
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
     mRNA            complement(join(36896. .39334,42146. .42257))
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
                     /product="corneodesmosin"
                     /note="match: ESTs: AU134804.1 BG676174.1
                     match: cDNAs: AF030130.1 BC031993.1 L20815.1"
     polyA_signal    complement(36915. .36920)
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
     polyA_site      complement(37383)
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
     polyA_signal    complement(37408. .37413)
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
     CDS             complement(join(37830. .39334,42146. .42230))
                     /gene="CDSN"
                     /locus_tag="DAQB-185I1.3-001"
                     /standard_name="OTTHUMP00000014908"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="CAI18479.1"
                     /db_xref="GI:56206477"
                     /db_xref="InterPro:IPR001917"
                     /db_xref="UniProt/TrEMBL:Q5SQ85"
                     /translation="MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSS
                     FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPC
                     SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
                     GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSP
```

```
                     GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
                     PSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGS
                     KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
                     LGPQLADPEVFLPQGELLDSP"
     gene            complement(join(59371. .60142,60807. .61185))
                     /gene="PSORS1C2"
                     /locus_tag="DAQB-185I1.2-001"
     mRNA            complement(join(59371. .60142,60807. .61185))
                     /gene="PSORS1C2"
                     /locus_tag="DAQB-185I1.2-001"
                     /product="psoriasis susceptibility 1 candidate 2"
                     /note="match: cDNAs: AB031480.1 AF159091.1"
     polyA_site      complement(59371)
                     /gene="PSORS1C2"
                     /locus_tag="DAQB-185I1.2-001"
     polyA_signal    complement(59387. .59392)
                     /gene="PSORS1C2"
                     /locus_tag="DAQB-185I1.2-001"
     CDS             complement(join(59787. .60142,60807. .60861))
                     /gene="PSORS1C2"
                     /locus_tag="DAQB-185I1.2-001"
                     /standard_name="OTTHUMP00000014715"
                     /note="match: proteins: Q9TNN6 Q9UIG4"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 2"
                     /protein_id="CAI18481.1"
                     /db_xref="GI:56206479"
                     /translation="MILNWKLLGILVLCLHTRGISGSEGHPSHPPAEDREEAGSPTLP
                     QGPPVPGDPWPGAPPLFEDPPPTRPSRPWRDLPETGVWPPEPPRTDPPQPPRPDDPWP
                     AGPQPPENPWPPAPEVDNRPQEEPDLDPPREEYR"
     gene            join(60195. .60295,61476. .61872)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-002"
     mRNA            join(60195. .60295,61476. .61872)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-002"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: BI825917.1"
     CDS             join(60592. .60614,61476. .61767)
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-002"
                     /standard_name="OTTHUMP00000044052"
                     /note="match: proteins: Q9UIG5 Sw:Q7YR46"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 1"
                     /protein_id="CAI18480.1"
                     /db_xref="GI:56206478"
                     /db_xref="UniProt/TrEMBL:Q5SQ84"
                     /translation="MEPANHFWHAGDLQAMISKEFHLAATQDDCRKGRTQEDILVPSS
                     HPELFASVLPMAPEEAARLQQPQPLPPPSGIHLSASRTLALTLLYSSPPSHSPFGLSS
                     LI"
     polyA_signal    61852. .61857
                     /gene="PSORS1C1"
                     /locus_tag="DAQB-185I1.5-002"

  Query Match             78.8%;  Score 1253;  DB 8;  Length 113388;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39337 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 39278

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39277 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 39218

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39217 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 39158

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39157 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 39098

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      39097 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 39038

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      39037 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 38978

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38977 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 38918

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      38917 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 38858

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      38857 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 38798

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38797 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 38738

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38737 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 38678

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38677 GGGGTCAGAGACCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 38618

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38617 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 38558

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38557 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 38498

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      38497 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCACCAAAGAGA 38438

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38437 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 38378

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38377 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 38318

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38317 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 38258

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38257 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 38198

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38197 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 38138

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38137 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 38078

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38077 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 38018
```

```
Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      38017 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 37958

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37957 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 37898

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37897 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 37838

Qy       1583 GTCCATAA 1590
              ||||||||
Db      37837 GTCCATAA 37830


RESULT 31
AX336639
LOCUS       AX336639               2530 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 7148 from Patent WO0194629.
ACCESSION   AX336639
VERSION     AX336639.1  GI:18127358
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 7148 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2530
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             78.3%;  Score 1245;  DB 6;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514


RESULT 32
AX786041
LOCUS       AX786041               2530 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequence 11 from Patent WO03050311.
ACCESSION   AX786041
VERSION     AX786041.1  GI:32953659
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pras,E., Levy-Nissenbaum,E., Frydman,M. and Thederahn,T.
  TITLE     Methods and compositions for diagnosing and treating hypotrichosis
            simplex
  JOURNAL   Patent: WO 03050311-A 11 19-JUN-2003;
            Wella AG (DE)
FEATURES             Location/Qualifiers
     source          1. .2530
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           78.3%;  Score 1245;  DB 6;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514


RESULT 33
HUMSPROT
LOCUS       HUMSPROT               2530 bp    mRNA    linear   PRI 02-SEP-1994
DEFINITION  Human S protein mRNA, complete cds.
ACCESSION   L20815
VERSION     L20815.1  GI:414809
KEYWORDS    S gene; S protein; cell differentiation; keratin; keratinocyte
            differentiation; loricrin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 2530)
  AUTHORS   Zhou,Y. and Chaplin,D.D.
  TITLE     Identification in the HLA class I region of a gene expressed late
            in keratinocyte differentiation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 90 (20), 9470-9474 (1993)
   PUBMED   8415725
COMMENT     Original source text: Homo sapiens (library: lambda gt10 (from
            Clontech)) neonate foreskin cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .2530
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="6p21.3"
                     /cell_type="keratinocyte"
                     /tissue_type="foreskin"
                     /dev_stage="neonate"
                     /germline
                     /tissue_lib="lambda gt10 (from Clontech)"
     CDS             63. .1523
                     /standard_name="S protein"
                     /note="epidermal granular cell layer-specific"
                     /codon_start=1
                     /protein_id="AAA21321.1"
                     /db_xref="GI:414810"
                     /translation="MLALLLAGLLLPGTLAKSIGTFSDPCKDPTRITSPNDPCLTGKG
                     DSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYS
                     SGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSAL
                     PTNDNSYRGILNPSQPGQSSSSSQTSGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVS
                     HSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKS
                     YGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKY
                     FSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSS
                     GSPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLT
                     LTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRIS"
     intron          99^100
                     /number=1
     polyA_signal    2511. .2516
     polyA_site      2530
ORIGIN

  Query Match             78.3%;  Score 1245;  DB 8;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

RESULT 34
```

```
HSA238461
LOCUS       HSA238461               1324 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA B62-Cw*0304.
ACCESSION   AJ238461
VERSION     AJ238461.1  GI:6010124
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
            Kronke,M.
  TITLE     Corneodesmosin gene polymorphism demonstrates strong linkage
            disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL   Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED   10599883
REFERENCE   2  (bases 1 to 1324)
  AUTHORS   Jenisch,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
            University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1324
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B62-Cw*0304"
     CDS             <1. .1276
                     /note="allele: 01"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57266.1"
                     /db_xref="GI:6010125"
                     /db_xref="UniProt/TrEMBL:Q9UBH8"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             70.6%;  Score 1123;  DB 8;  Length 1324;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1273;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60

Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACTCGACAGTCCATAA 1590
              ||||||||||||||||
Db       1261 ACTCGACAGTCCATAA 1276


RESULT 35
HSA238462
LOCUS       HSA238462               1324 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA B62-Cw*0304.
ACCESSION   AJ238462
```

```
VERSION     AJ238462.1  GI:6010126
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
            Kronke,M.
  TITLE     Corneodesmosin gene polymorphism demonstrates strong linkage
            disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL   Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED   10599883
REFERENCE   2  (bases 1 to 1324)
  AUTHORS   Jenisch,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
            University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1324
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B62-Cw*0304"
     CDS             <1. .1276
                     /note="allele: 01"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57267.1"
                     /db_xref="GI:6010127"
                     /db_xref="UniProt/TrEMBL:Q9UBH8"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTSGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             67.4%;  Score 1072;  DB 8;  Length 1324;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1272;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60

Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTATGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACTCGACAGTCCATAA 1590
              ||||||||||||||||
Db       1261 ACTCGACAGTCCATAA 1276


RESULT 36
HSA238463
LOCUS       HSA238463               1324 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA
            B57-Cw*0602,B37-Cw*0602.
ACCESSION   AJ238463
VERSION     AJ238463.1  GI:6010128
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
            Kronke,M.
  TITLE     Corneodesmosin gene polymorphism demonstrates strong linkage
            disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL   Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED   10599883
REFERENCE   2  (bases 1 to 1324)
  AUTHORS   Jenisch,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
            University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1324
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B57-Cw*0602,B37-Cw*0602"
     CDS             <1. .1276
                     /note="allele: 02"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57268.1"
                     /db_xref="GI:6010129"
                     /db_xref="UniProt/TrEMBL:Q9UIP0"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             67.4%;  Score 1072;  DB 8;  Length 1324;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1272;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60

Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACTCGACAGTCCATAA 1590
              ||||||||||||||||
Db        1261 ACTCGACAGTCCATAA 1276


RESULT 37
HSA238464
LOCUS       HSA238464               1324 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA B13-Cw*0602.
ACCESSION   AJ238464
VERSION     AJ238464.1  GI:6010130
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
            Kronke,M.
  TITLE     Corneodesmosin gene polymorphism demonstrates strong linkage
            disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL   Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED   10599883
REFERENCE   2  (bases 1 to 1324)
  AUTHORS   Jenisch,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
            University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1324
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B13-Cw*0602"
     CDS             <1. .1276
                     /note="allele: 03"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57269.1"
                     /db_xref="GI:6010131"
                     /db_xref="GOA:Q9NRH7"
                     /db_xref="UniProt/TrEMBL:Q9NRH7"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSGSSISSSSGSPYHPCGSASQSPCSPPGTGSFSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             64.2%;  Score 1021;  DB 8;  Length 1324;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1271;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60

Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCGGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACTCGACAGTCCATAA 1590
              ||||||||||||||||
Db       1261 ACTCGACAGTCCATAA 1276


RESULT 38
HSA238465
LOCUS       HSA238465               1324 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA B18-Cw*0701.
ACCESSION   AJ238465
VERSION     AJ238465.1  GI:6010132
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
```

```
          Kronke,M.
  TITLE   Corneodesmosin gene polymorphism demonstrates strong linkage
          disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED 10599883
REFERENCE 2  (bases 1 to 1324)
  AUTHORS Jenisch,S.
  TITLE   Direct Submission
  JOURNAL Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
          University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1324
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B18-Cw*0701"
     CDS             <1. .1276
                     /note="allele: 04"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57270.1"
                     /db_xref="GI:6010133"
                     /db_xref="GOA:Q9NPG5"
                     /db_xref="UniProt/TrEMBL:Q9NPG5"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSSSSISSSAGSPYHPCGSASQSPCSPPGTGSFSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             64.2%;  Score 1021;  DB 8;  Length 1324;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1271;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60

Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGACCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        481 TGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACTCGACAGTCCATAA 1590
              ||||||||||||||||
Db       1261 ACTCGACAGTCCATAA 1276


RESULT 39
AF224750
LOCUS       AF224750               1761 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-1.32 allele, partial
            cds.
ACCESSION   AF224750
VERSION     AF224750.1  GI:8699029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1761)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
```

```
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1761)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1761
                     /gene="CDSN"
                     /allele="CDSN-1.32"
     mRNA            <108. .>1761
                     /gene="CDSN"
                     /allele="CDSN-1.32"
                     /product="corneodesmosin"
     CDS             <108. .1607
                     /gene="CDSN"
                     /allele="CDSN-1.32"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78562.1"
                     /db_xref="GI:8699030"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCFTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSY
                     LVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAAS
                     SAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGSPYHPCGSASQSP
                     CSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS
                     AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match           63.5%;  Score 1010;  DB 8;  Length 1761;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1500;  Conservative    0;  Mismatches    5;  Indels    3;  Gaps    1;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        163 TCACCTCCCCTAACGACCCCTGCTTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 462

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 522

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        523 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 582

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        583 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTTTGGGGTATCCAGCAGTG 642

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 702

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 762

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        763 GGGGTCAGAGGCCT---GTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 819

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 879

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 939

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 999

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1059

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1119

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1179

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1239

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1299

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1359

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1419

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1479

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1480 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1539

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1599

Qy       1583 GTCCATAA 1590
              ||||||||
Db       1600 GTCCATAA 1607
```

```
RESULT 40
CS086650
LOCUS       CS086650               2284 bp    DNA     linear   PAT 25-MAY-2005
DEFINITION  Sequence 12 from Patent EP1531161.
ACCESSION   CS086650
VERSION     CS086650.1  GI:66712119
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Lal,P., Corley,N.C., Guegler,K.J., Patterson,C.,
            Baughn,M.R. and Yue,H.
  TITLE     Human epidermal proteins HEPI-1 to HEPI-6
  JOURNAL   Patent: EP 1531161-A 12 18-MAY-2005;
            Incyte Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2284
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             61.5%;  Score 978;  DB 6;  Length 2284;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1578;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||
Db       1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||
Db       1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636


RESULT 41
BA000041_03
WPCOMMENT
  Sequence split into 18 fragments  LOCUS BA000041 Accession BA000041
     Fragment Name      Begin     End
     BA000041_00            1    110000
```

```
     BA000041_01        100001    210000
     BA000041_02        200001    310000
     BA000041_03        300001    410000
     BA000041_04        400001    510000
     BA000041_05        500001    610000
     BA000041_06        600001    710000
     BA000041_07        700001    810000
     BA000041_08        800001    910000
     BA000041_09        900001   1010000
     BA000041_10       1000001   1110000
     BA000041_11       1100001   1210000
     BA000041_12       1200001   1310000
     BA000041_13       1300001   1410000
     BA000041_14       1400001   1510000
     BA000041_15       1500001   1610000
     BA000041_16       1600001   1710000
     BA000041_17       1700001   1750601
Continuation (4 of 18) of BA000041 from base 300001 (BA000041 Pan troglodytes DNA, major

  Query Match           56.4%;  Score 896;  DB 8;  Length 110000;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1496;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      95545 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCATGCGTA 95604

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95605 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 95664

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95665 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 95724

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95725 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 95784

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      95785 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 95844

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db      95845 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCAGGAAGCAGCGGCTCTCACT 95904

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      95905 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 95964

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      95965 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 96024

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db      96025 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACCTTTGGGGTATCCAGCAGTG 96084

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96085 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 96144

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      96145 GTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 96204

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      96205 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCATGGTTCTGGTGCCCCTGGAGTGG 96264

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      96265 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 96324

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96325 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 96384

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      96385 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCACCAAAGAGA 96444

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96445 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 96504

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96505 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 96564

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96565 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 96624

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96625 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 96684

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96685 CCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 96744

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      96745 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGCAAAATCATCCTTCAGCCTT 96804

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96805 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 96864

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96865 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 96924

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96925 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 96984

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96985 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 97044

Qy       1583 GTCCATAA 1590
              ||||||||
Db      97045 GTCCATAA 97052
```

```
RESULT 42
AF224757
LOCUS       AF224757               1767 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-2.22 allele, partial
            cds.
ACCESSION   AF224757
VERSION     AF224757.1  GI:8699043
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1767)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
```

```
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1767)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1767
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1767
                     /gene="CDSN"
                     /allele="CDSN-2.22"
     mRNA            <108. .>1767
                     /gene="CDSN"
                     /allele="CDSN-2.22"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-2.22"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78569.1"
                     /db_xref="GI:8699044"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGNSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLNSP"
ORIGIN

  Query Match             56.2%;  Score 894;  DB 8;  Length 1767;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1444;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 203

Qy        184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TCCAGCGGTTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 263

Qy        244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 323

Qy        304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 383

Qy        364 GGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 423
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 443

Qy        424 GGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 483
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        444 GGAAGCAGCGGCTCTCACTCGGGAAACAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 503

Qy        484 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 563

Qy        544 TCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACC 603
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||

Db        564 TCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACC 623

Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 TTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 683

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 743

Qy        724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 803

Qy        784 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 863

Qy        844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 923

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        924 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTG 983

Qy        964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1043

Qy       1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1103

Qy       1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1104 GCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1163

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1164 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1223

Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263
              |||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||||
Db       1224 TCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1283

Qy       1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1284 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGC 1343

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1344 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1403

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1463

Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1523

Qy       1504 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1583

Qy       1564 CAAGGAGAGTTACTC 1578
              |||||||||||||||
Db       1584 CAAGGAGAGTTACTC 1598


RESULT 43
AF224758
```

```
LOCUS       AF224758               1767 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-2.23 allele, partial
            cds.
ACCESSION   AF224758
VERSION     AF224758.1  GI:8699045
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1767)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1767)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1767
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1767
                     /gene="CDSN"
                     /allele="CDSN-2.23"
     mRNA            <108. .>1767
                     /gene="CDSN"
                     /allele="CDSN-2.23"
                     /product="corneodesmosin"
     CDS             <108. .1610
                     /gene="CDSN"
                     /allele="CDSN-2.23"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78570.1"
                     /db_xref="GI:8699046"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSSGSHSGNSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNP
                     SQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVS
                     GGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDS
                     YLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAA
                     SSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQS
                     PCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDP
                     SAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVPLPQGELLNSP"
ORIGIN

  Query Match             56.2%;  Score 894;  DB 8;  Length 1767;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1444;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 203

Qy        184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 263

Qy        244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 323

Qy        304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        324 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 383

Qy        364 GGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 423
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 443

Qy        424 GGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 483
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        444 GGAAGCAGCGGCTCTCACTCGGGAAACAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 503

Qy        484 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 563

Qy        544 TCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACC 603
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        564 TCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACC 623

Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 TTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 683

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        684 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCC 743

Qy        724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 803

Qy        784 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 863

Qy        844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 923

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        924 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTG 983

Qy        964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1043

Qy       1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1103

Qy       1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1104 GCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1163

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1164 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1223

Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263
              |||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db       1224 TCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1283

Qy       1264 CCCTGCTCCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1284 CCCTGCTCCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGC 1343

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1344 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1403

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1463
```

```
Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1523

Qy       1504 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1583

Qy       1564 CAAGGAGAGTTACTC 1578
             |||||||||||||||
Db       1584 CAAGGAGAGTTACTC 1598
```

RESULT 44
AF224756
LOCUS AF224756 1767 bp DNA linear PRI 14-MAR-2001
DEFINITION Homo sapiens corneodesmosin (CDSN) gene, CDSN-2.21 allele, partial cds.
ACCESSION AF224756
VERSION AF224756.1 GI:8699041
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1767)
AUTHORS Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and Serre,G.
TITLE Identification of six novel polymorphisms in the human corneodesmosin gene
JOURNAL Tissue Antigens 57 (1), 32-38 (2001)
PUBMED 11169256
REFERENCE 2 (bases 1 to 1767)
AUTHORS Guerrin-Weber,M. and Vincent,C.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37, allees Jules-Guesde, Toulouse 31073, France
FEATURES Location/Qualifiers
source 1. .1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.31"
gene <108. .>1767
/gene="CDSN"
/allele="CDSN-2.21"
mRNA <108. .>1767
/gene="CDSN"
/allele="CDSN-2.21"
/product="corneodesmosin"
CDS <108. .1610
/gene="CDSN"
/allele="CDSN-2.21"
/codon_start=1
/product="corneodesmosin"
/protein_id="AAF78568.1"
/db_xref="GI:8699042"
/translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGNSSSHSSSSSSFQPSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLNSP"
ORIGIN

Query Match 53.0%; Score 843; DB 8; Length 1767;
Best Local Similarity 99.2%; Pred. No. 0;

Matches 1443; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 203

Qy        184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TCCAGCGGTTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 263

Qy        244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 323

Qy        304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 383

Qy        364 GGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 423
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 443

Qy        424 GGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 483
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        444 GGAAGCAGCGGCTCTCACTCGGGAAACAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 503

Qy        484 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 563

Qy        544 TCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACC 603
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        564 TCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACC 623

Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 TTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 683

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        684 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCC 743

Qy        724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 803

Qy        784 TCTGGTGCCCCCTGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 TCTGGTGCCCCCTGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 863

Qy        844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 923

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        924 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTG 983

Qy        964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1043

Qy       1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1103

Qy       1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1104 GCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1163

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1164 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1223
```

```
Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263
             ||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||||
Db       1224 TCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1283

Qy       1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1284 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGC 1343

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1344 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1403

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1463

Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1523

Qy       1504 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1583

Qy       1564 CAAGGAGAGTTACTC 1578
             |||||||||||||||
Db       1584 CAAGGAGAGTTACTC 1598
```

RESULT 45
AL662844/c

```
LOCUS       AL662844              204298 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-299F13 on chromosome 6 contains
            the gene for STG protein, the CDSN gene for corneodesmosin, the
            gene for SEEK1 protein, the gene for SPR1 protein, the gene for
            HCR, the TCF19 gene for transcription factor 19, the POU5F1 gene
            for POU domain, class 5, transcription factor 1, a putative novel
            transcript and four CpG islands, complete sequence.
ACCESSION   AL662844
VERSION     AL662844.12  GI:21212039
KEYWORDS    HTG; a-helix coiled-coiled rod homologue; CDSN; corneodesmosin; CpG
            island; HCR; OCT3; Oct4; OTF4; OTF5; POU domain, class 5,
            transcription factor 1; POU5F1; PSORS1C3; S; SC1; SEEK1; SPR1; STG;
            TCF19; transcription factor 19.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 204298)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 25, 2002 this sequence version replaced gi:20330233.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs constructed by
            the MHC Haplotype Consortium and collaborators. Further information
            can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            XXbac-299F13 is from a CHORI-501 human bac - PGF cell line library
            VECTOR: pTARBAC2.1
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .204298
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-299F13"
                     /clone_lib="CHORI-501"
     misc_feature    2000
                     /note="Clone_right_end: XXbac-118E17"
     gene            complement(50711. .52043)
                     /gene="C6orf15"
                     /locus_tag="XXbac-BPG299F13.10-001"
     mRNA            complement(join(50711. .51779,51977. .52043))
                     /gene="C6orf15"
                     /locus_tag="XXbac-BPG299F13.10-001"
                     /product="chromosome 6 open reading frame 15"
                     /note="match: cDNAs: AB031481.1"
     CDS             complement(join(50869. .51779,51977. .52043))
                     /gene="C6orf15"
                     /locus_tag="XXbac-BPG299F13.10-001"
                     /standard_name="OTTHUMP00000029146"
                     /note="match: proteins: Q9BGL9"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 15"
                     /protein_id="CAI18322.1"
                     /db_xref="GI:55962038"
                     /db_xref="UniProt/TrEMBL:Q86Z05"
                     /translation="MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNLGTNL
                     PQLGQPSSTGPSNSEHPQPALDPRSNDLARVPLKLSVPASDGFPPAGGSAVQRWPPSW
                     GLPAMDSWPPEDPWQMMAAAAEDRLGEALPEELSYLSSAAALAPGSGPLPGESSPDAT
                     GLSPKASLLHQDSESRRLPRSNSLGAGGKILSQRPPWSLIHRVLPDHPWGTLNPSVSW
                     GGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGSWGNINRYPGGSWGNIN
                     RYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG"
     gene            join(54312. .54379,65164. .65327,69057. .69133,77918. .77947,
                     78144. .78267,79129. .79580)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-001"
     mRNA            join(54312. .54379,65164. .65327,69057. .69133,77918. .77947,
                     78144. .78267,79129. .79580)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-001"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: cDNAs: AB031479.1 Em:AF484418.1"
     gene            join(54319. .54379,79129. .79526)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-003"
     mRNA            join(54319. .54379,79129. .79526)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-003"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: Em:AI343394.1
                     match: cDNAs: Em:AF484419.1"
     gene            complement(join(54581. .57017,59823. .59934))
                     /gene="CDSN"
                     /locus_tag="XXbac-BPG299F13.1-001"
     mRNA            complement(join(54581. .57017,59823. .59934))
                     /gene="CDSN"
                     /locus_tag="XXbac-BPG299F13.1-001"
                     /product="corneodesmosin"
                     /note="match: ESTs: AI746833.1 AU134804.1 BG676174.1
                     match: cDNAs: AF030130.1 L20815.1"
     CDS             complement(join(55513. .57017,59823. .59907))
                     /gene="CDSN"
                     /locus_tag="XXbac-BPG299F13.1-001"
```

```
                     /standard_name="OTTHUMP00000029207"
                     /note="match: proteins: O43509 Q15517 Q95IF9 Q9UIN6"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="CAI18323.1"
                     /db_xref="GI:55962039"
                     /db_xref="InterPro:IPR001917"
                     /db_xref="UniProt/TrEMBL:Q5STD2"
                     /translation="MGSSRAPWMGRVGGHGMMALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGNSGSHSGSSSSHSSSSSS
                     FQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPC
                     SSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNG
                     GLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSP
                     GVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCS
                     PSSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSQSSGKIILQPCGS
                     KSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKP
                     LGPQLADPEVFLPQGELLNSP"
CDS                  join(69121. .69133,77918. .77947,78144. .78267,79129. .79420)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-003"
                     /standard_name="OTTHUMP00000029096"
                     /note="match: proteins: Q96QC3 Q9UIG5"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 1"
                     /protein_id="CAI18324.1"
                     /db_xref="GI:55962040"
                     /db_xref="UniProt/TrEMBL:Q5ST21"
                     /translation="MTCTDQKSHSQRALGTQTPALQGPQLLNTDPSSEETRPPHVNPD
                     RLCHMEPANHFWHAGDLQAMISKEFHLAATQDDCRKGRTQEDILVPSSHPELFASVLP
                     MAPEEAARLQQPQPLPPPSGIHLSASRTLAPTLLYSSPPSHSPFGLSSLI"
gene                 complement(join(77022. .77794,78460. .78829))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BPG299F13.2-001"
mRNA                 complement(join(77022. .77794,78460. .78829))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BPG299F13.2-001"
                     /product="psoriasis susceptibility 1 candidate 2"
                     /note="match: cDNAs: AB031480.1"
CDS                  complement(join(77439. .77794,78460. .78514))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BPG299F13.2-001"
                     /standard_name="OTTHUMP00000029097"
                     /note="match: proteins: Q9UIG4"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 2"
                     /protein_id="CAI18325.1"
                     /db_xref="GI:55962041"
                     /db_xref="UniProt/TrEMBL:Q5STD0"
                     /translation="MILNWKLLGILVLCLHTRGISGSEGHPSHPPAEDREEAGSPTLP
                     QGPPVPGDPWPGAPPLFEDPPPTRPSRPWRDLPETGVWLPEPPRTDPPQPPRPDDPWP
                     AGPQPPENPWPPAPEVDNRPQEEPDLDPPREEYR"
gene                 join(77845. .77947,79129. .79526)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-002"
mRNA                 join(77845. .77947,79129. .79526)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BPG299F13.4-002"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: BI825917.1"
gene                 80218. .80401
                     /locus_tag="XXbac-BPG299F13.6-001"
                     /pseudo
CDS                  80218. .80401
                     /locus_tag="XXbac-BPG299F13.6-001"
                     /pseudo
                     /codon_start=1
                     /product="polymerase (RNA) II (DNA directed) polypeptide
                     L, 7.6kDa (POLR2L) pseudogene"
gene                 complement(join(81927. .82204,82451. .82622,82750. .82901,
                     84175. .84276,84373. .84561,84654. .84790,84936. .84981,
                     85184. .85296,87893. .87999,88106. .88216,89554. .89703,
                     89943. .90053,90213. .90348,90447. .90610,93984. .94287,
                     96219. .96432,96511. .96577,97655. .97693))
                     /gene="C6orf18"
                     /locus_tag="XXbac-BPG299F13.11-001"
mRNA                 complement(join(81927. .82204,82451. .82622,82750. .82901,
                     84175. .84276,84373. .84561,84654. .84790,84936. .84981,
                     85184. .85296,87893. .87999,88106. .88216,89554. .89703,
                     89943. .90053,90213. .90348,90447. .90610,93984. .94287,
                     96219. .96432,96511. .96577,97655. .97693))
                     /gene="C6orf18"
                     /locus_tag="XXbac-BPG299F13.11-001"
                     /product="chromosome 6 open reading frame 18"
                     /note="match: ESTs: BI820818.1
                     match: cDNAs: AB029331.1 AF216493.1 AK000204.1 AK000217.1

  Query Match             53.0%;  Score 843;  DB 8;  Length 204298;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1443;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
              ||||||
```

```
Qy        964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56139 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 56080

Qy       1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56079 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 56020

Qy       1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      56019 GCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 55960

Qy       1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55959 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 55900

Qy       1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263
              |||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||||
Db      55899 TCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 55840

Qy       1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      55839 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGC 55780

Qy       1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55779 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 55720

Qy       1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55719 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 55660

Qy       1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55659 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 55600

Qy       1504 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55599 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 55540

Qy       1564 CAAGGAGAGTTACTC 1578
              |||||||||||||||
Db      55539 CAAGGAGAGTTACTC 55525


RESULT 46
HSM807103/c
LOCUS       HSM807103               3889 bp    mRNA    linear   PRI 28-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686G23132 (from clone DKFZp686G23132).
ACCESSION   BX640954
VERSION     BX640954.1  GI:34365293
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3889)
  AUTHORS   Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp686G23132) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .3889
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="6"
                     /clone="DKFZp686G23132"
                     /tissue_type="human uterus"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_signal    3842. .3847
     polyA_site      3865
ORIGIN

  Query Match             49.8%;  Score 792;  DB 8;  Length 3889;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1442;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2498 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 2439

Qy        184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2438 TCCAGCGGTTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 2379

Qy        244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2378 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 2319

Qy        304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2318 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 2259

Qy        364 GGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 423
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2258 GGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCG 2199

Qy        424 GGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 483
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2198 GGAAACAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTT 2139

Qy        484 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2138 CAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAAC 2079

Qy        544 TCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACC 603
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       2078 TCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACC 2019

Qy        604 TCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 663
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2018 TTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGAC 1959

Qy        664 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCC 723
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1958 ATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCC 1899

Qy        724 AGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 783
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1898 AGCTCCCACTCTGTCTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGT 1839

Qy        784 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1838 TCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAG 1779

Qy        844 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1778 CCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGC 1719

Qy        904 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTG 963
```

```
Db      1718 TCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTG 1659

Qy       964 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1023

Db      1658 GGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCT 1599

Qy      1024 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1083

Db      1598 GGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCG 1539

Qy      1084 GCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1143

Db      1538 GCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGA 1479

Qy      1144 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1203

Db      1478 GGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGT 1419

Qy      1204 TCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1263

Db      1418 TCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGC 1359

Qy      1264 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGC 1323

Db      1358 CCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGC 1299

Qy      1324 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1383

Db      1298 AAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCT 1239

Qy      1384 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1443

Db      1238 GTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCT 1179

Qy      1444 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1503

Db      1178 GGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGAT 1119

Qy      1504 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1563

Db      1118 ATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCC 1059

Qy      1564 CAAGGAGAGTTACTC 1578

Db      1058 CAAGGAGAGTTACTC 1044
```

RESULT 47
BT019904
LOCUS BT019904 1587 bp mRNA linear SYN 28-OCT-2004
DEFINITION Synthetic construct Homo sapiens corneodesmosin mRNA, partial cds.
ACCESSION BT019904
VERSION BT019904.1 GI:54696669
KEYWORDS FLI_CDNA.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1587)
AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1587)
AUTHORS Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-2004) BD Biosciences Clontech, 1020 East Meadow circle, Palo Alto, California 94303, USA
COMMENT This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES Location/Qualifiers
source 1. .1587
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH01206L1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
CDS 1. .>1587
/note="Mutations: 52:M->L;447:3DEL;597:OK"
/codon_start=1
/transl_table=11
/product="corneodesmosin"
/protein_id="AAV38707.1"
/db_xref="GI:54696670"
/translation="MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKD
PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSHSSSSSSF
QFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCS
SDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGG
LPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG
VPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSP
SSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSK
SSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL
GPQLADPEVFLPQGELLDSPL"
ORIGIN

Query Match 46.1%; Score 733; DB 11; Length 1587;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1573; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

```
Qy         1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Db         1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 60

Qy        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Db        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Db       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Db       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Db       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Db       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Db       361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Db       421 TCGGGAAGCAGCGGCTCTCACTCGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 477
```

```
Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db       478 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 537

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db       538 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 597

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db       598 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 657

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db       658 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 717

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db       718 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 777

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
Db       778 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 837

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
Db       838 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 897

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
Db       898 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 957

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
Db       958 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1017

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db      1018 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1077

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db      1078 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1137

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db      1138 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1197

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db      1198 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1257

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db      1258 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1317

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db      1318 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1377

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db      1378 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1437

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db      1438 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1497

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
Db      1498 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1557
```

```
Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCAT 1588
Db      1558 CCCCAAGGAGAGTTACTCGACAGTCCAT 1585
```

RESULT 48
AF224755

```
LOCUS       AF224755               1764 bp    DNA     linear   PRI 14-MAR-2001
DEFINITION  Homo sapiens corneodesmosin (CDSN) gene, CDSN-2.11 allele, partial
            cds.
ACCESSION   AF224755
VERSION     AF224755.1  GI:8699039
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1764)
  AUTHORS   Guerrin,M., Vincent,C., Simon,M., Tazi Ahnini,R., Fort,M. and
            Serre,G.
  TITLE     Identification of six novel polymorphisms in the human
            corneodesmosin gene
  JOURNAL   Tissue Antigens 57 (1), 32-38 (2001)
   PUBMED   11169256
REFERENCE   2  (bases 1 to 1764)
  AUTHORS   Guerrin-Weber,M. and Vincent,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) LBPDM, Faculte de Medecine Purpan, 37,
            allees Jules-Guesde, Toulouse 31073, France
FEATURES             Location/Qualifiers
     source          1. .1764
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.31"
     gene            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-2.11"
     mRNA            <108. .>1764
                     /gene="CDSN"
                     /allele="CDSN-2.11"
                     /product="corneodesmosin"
     CDS             <108. .1607
                     /gene="CDSN"
                     /allele="CDSN-2.11"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="AAF78567.1"
                     /db_xref="GI:8699040"
                     /translation="TLAKSIGTFSDPCKDPTRITSPNDPCLTGKGDSSGFSSYSGSSS
                     SGSSISSARSSGGGSSGSSSGSSIAQGGSAGSFKPGTGYSQVSYSSGSGSSLQGASGS
                     SQLGSSSSHSGSGSHSGSSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPS
                     QPGQSSSSSQTFGVSSSGQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSG
                     GQRPVVVVVDQHGSGAPGVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSY
                     LVPGMTYSKGKIYPVGYFTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAAS
                     SAIAFQPVGTGGVQLCGGGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQSP
                     CSPPGTGSFSSSSSSQSSGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPS
                     AGAKPCGSSSAGKIPCRSIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN
```

```
  Query Match             44.3%;  Score 704;  DB 8;  Length 1764;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1494;  Conservative    0;  Mismatches   11;  Indels    3;  Gaps    1;

Qy        83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
Db       103 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 162

Qy       143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
Db       163 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 222
```

```
Qy       203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       223 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 282

Qy       263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       283 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 342

Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       343 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 402

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       403 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGG---GCAGCGGCTCTCACT 459

Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       460 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 519

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       520 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 579

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
            |||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db       580 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACCTTTGGGGTATCCAGCAGTG 639

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       640 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 699

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
            |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       700 GTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 759

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       760 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 819

Qy       803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       820 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 879

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       880 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 939

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       940 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCACCAAAGAGA 999

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1000 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1059

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1060 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1119

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1120 CATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1179

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1180 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 1239

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1240 CCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1299

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      1300 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGCAAAATCATCCTTCAGCCTT 1359

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
            | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1360 GCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1419

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1420 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1479

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1480 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1539

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1540 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1599

Qy      1583 GTCCATAA 1590
            ||||||||
Db      1600 GTCCATAA 1607
```

RESULT 49
AL662867/c

```
LOCUS       AL662867              70288 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone XXbac-34D1 on chromosome 6 contains
            the genes for STG protein, SEEK1 protein, SPR1 protein and the CDSN
            gene for corneodesmosin, complete sequence.
ACCESSION   AL662867
VERSION     AL662867.5  GI:20068664
KEYWORDS    HTG; CDSN; corneodesmosin; S; SEEK1; SPR1; STG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 70288)
  AUTHORS   Sycamore,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 7, 2002 this sequence version replaced gi:19309560.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep XXbac-34D1 is
            from a CHORI-502 human bac - COX cell line library VECTOR:
            pTARBAC2.1
            This sequence was generated from part of bacterial clone contigs
            constructed by the MHC Haplotype Consortium and collaborators.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6/MHC
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .70288
                     /organism="Homo sapiens"
```

```
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="XXbac-34D1"
                     /clone_lib="CHORI-502"
     misc_feature    2000
                     /note="Clone_right_end: XXbac-31G15"
     gene            complement(40621. .41953)
                     /gene="C6orf15"
                     /locus_tag="XXbac-BCX34D1.4-001"
     mRNA            complement(join(40621. .41689,41887. .41953))
                     /gene="C6orf15"
                     /locus_tag="XXbac-BCX34D1.4-001"
                     /product="chromosome 6 open reading frame 15"
                     /note="match: cDNAs: AB031481.1"
     CDS             complement(join(40779. .41689,41887. .41953))
                     /gene="C6orf15"
                     /locus_tag="XXbac-BCX34D1.4-001"
                     /standard_name="OTTHUMP00000029556"
                     /note="match: proteins: Q9BGL9"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 15"
                     /protein_id="CAI17718.1"
                     /db_xref="GI:55961419"
                     /db_xref="UniProt/TrEMBL:Q6UXA7"
                     /translation="MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNL
                     PQLGQPSSTGPSNSEHPQPALDPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSW
                     GLPAMDSWPPEDPWQMMAAAAEDRLGEALPEELSYLSSAAALAPGSGPLPGESSPDAT
                     GLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRPPWSLIHRVLPDHPWGTLNPSVSW
                     GGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGSWGNINRYPGGSWGNIN
                     RYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG"
     gene            join(44222. .44289,55077. .55240,58982. .59058,67846. .67875,
                     68072. .68195,69057. .69508)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-001"
     mRNA            join(44222. .44289,55077. .55240,58982. .59058,67846. .67875,
                     68072. .68195,69057. .69508)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-001"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: cDNAs: AB031479.1"
     gene            join(44229. .44289,69057. .69454)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-003"
     mRNA            join(44229. .44289,69057. .69454)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-003"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: Em:AI343394.1
                     match: cDNAs: Em:AF484419.1"
     gene            complement(join(44491. .46924,49734. .49832))
                     /gene="CDSN"
                     /locus_tag="XXbac-BCX34D1.1-001"
     mRNA            complement(join(44491. .46924,49734. .49832))
                     /gene="CDSN"
                     /locus_tag="XXbac-BCX34D1.1-001"
                     /product="corneodesmosin"
                     /note="match: cDNAs: AF030130.1 L20815.1"
     CDS             complement(join(45423. .46924,49734. .49818))
                     /gene="CDSN"
                     /locus_tag="XXbac-BCX34D1.1-001"
                     /standard_name="OTTHUMP00000029619"
                     /note="match: proteins: O43509 Q95IF9 Q9NP52 Q9NPE0 Q9NRH4
                     Q9NRH6 Q9NRH7"
                     /codon_start=1
                     /product="corneodesmosin"
                     /protein_id="CAI17719.1"
                     /db_xref="GI:55961420"
                     /db_xref="InterPro:IPR001917"
                     /db_xref="UniProt/TrEMBL:Q8IZU6"
                     /translation="MGSSRAPWMGRVGGHGMLALLLAGLLLPGTLAKSIGTFSDPCKD
                     PTRITSPNDPCLTGKGDSSGFSSYSGSSSSGSSISSARSSGGGSSGSSSGSSIAQGGS
                     AGSFKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGSSGSHSGSSSHSSSSSSF
                     QFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSSGQSVSSNQRPCS
                     SDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAPGVVQGPPCSNGG
                     LPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGYFTKENPVKGSPG
                     VPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCGGGSTGSKGPCSP
                     SSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQSSGKIILQPCGSK
                     SSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCRSIRDILAQVKPL
                     GPQLADPEVFLPQGELLDSP"
     CDS             join(59046. .59058,67846. .67875,68072. .68195,69057. .69348)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-003"
                     /standard_name="OTTHUMP00000029504"
                     /note="match: proteins: Q96QC3 Q9UIG5"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 1"
                     /protein_id="CAI17720.1"
                     /db_xref="GI:55961421"
                     /db_xref="UniProt/TrEMBL:Q5ST21"
                     /translation="MTCTDQKSHSQRALGTQTPALQGPQLLNTDPSSEETRPPHVNPD
                     RLCHMEPANHFWHAGDLQAMISKEFHLAATQDDCRKGRTQEDILVPSSHPELFASVLP
                     MAPEEAARLQQPQPLPPPSGIHLSASRTLAPTLLYSSPPSHSPFGLSSLI"
     gene            complement(join(66950. .67722,68388. .68757))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BCX34D1.2-001"
     mRNA            complement(join(66950. .67722,68388. .68757))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BCX34D1.2-001"
                     /product="psoriasis susceptibility 1 candidate 2"
                     /note="match: cDNAs: AB031480.1 AF159091.1"
     CDS             complement(join(67367. .67722,68388. .68442))
                     /gene="PSORS1C2"
                     /locus_tag="XXbac-BCX34D1.2-001"
                     /standard_name="OTTHUMP00000029505"
                     /note="match: proteins: Q9TNN6 Q9UIG4"
                     /codon_start=1
                     /product="psoriasis susceptibility 1 candidate 2"
                     /protein_id="CAI17721.1"
                     /db_xref="GI:55961422"
                     /translation="MILNWKLLGILVLCLHTRGISGSEGHPSHPPAEDREEAGSPTLP
                     QGPPVPGDPWPGAPPLFEDPPPTRPSRPWRDLPETGVWPPEPPRTDPPQPPRPDDPWP
                     AGPQPPENPWPPAPEVDNRPQEEPDLDPPREEYR"
     gene            join(67773. .67875,69057. .69454)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-002"
     mRNA            join(67773. .67875,69057. .69454)
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-002"
                     /product="psoriasis susceptibility 1 candidate 1"
                     /note="match: ESTs: BI825917.1"
     misc_feature    68289
                     /gene="PSORS1C1"
                     /locus_tag="XXbac-BCX34D1.3-002"
                     /note="Clone_left_end: XXbac-101P6"
ORIGIN

  Query Match           44.3%;  Score 704;  DB 8;  Length 70288;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1494;  Conservative    0;  Mismatches   11;  Indels    3;  Gaps    1;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46927 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 46868

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46867 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 46808

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46807 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 46748

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46747 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 46688
```

```
Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      46687 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 46628

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46627 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 46568

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46567 CGG---GCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 46511

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      46510 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 46451

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              |||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db      46450 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACCTTTGGGGTATCCAGCAGTG 46391

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46390 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 46331

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      46330 GTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 46271

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46270 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 46211

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46210 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 46151

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46150 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 46091

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      46090 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCACCAAAGAGA 46031

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46030 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 45971

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45970 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 45911

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45910 CATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 45851

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45850 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 45791

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45790 CCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 45731

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      45730 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGCAAAATCATCCTTCAGCCTT 45671

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45670 GCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 45611

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45610 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 45551

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45550 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 45491

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      45490 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 45431

Qy       1583 GTCCATAA 1590
              ||||||||
Db      45430 GTCCATAA 45423
```

RESULT 50
HSA238466

```
LOCUS       HSA238466               1327 bp    DNA     linear   PRI 27-JUL-2001
DEFINITION  Homo sapiens CDSN gene, partial, haplotype HLA B7-Cw*0702.
ACCESSION   AJ238466
VERSION     AJ238466.1  GI:6010134
KEYWORDS    CDSN gene; corneodesmosin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jenisch,S., Koch,S., Henseler,T., Nair,R.P., Elder,J.T.,
            Watts,C.E., Westphal,E., Voorhees,J.J., Christophers,E. and
            Kronke,M.
  TITLE     Corneodesmosin gene polymorphism demonstrates strong linkage
            disequilibrium with HLA and association with psoriasis vulgaris
  JOURNAL   Tissue Antigens 54 (5), 439-449 (1999)
   PUBMED   10599883
REFERENCE   2  (bases 1 to 1327)
  AUTHORS   Jenisch,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Jenisch S., Institute of Immunology,
            University of Kiel, Michaelisstrasse 5, Kiel 24105, Germany
FEATURES             Location/Qualifiers
     source          1. .1327
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="6p21"
                     /haplotype="HLA B7-Cw*0702"
     CDS             <1. .1276
                     /note="allele: 05"
                     /codon_start=2
                     /product="corneodesmosin"
                     /protein_id="CAB57271.1"
                     /db_xref="GI:6010135"
                     /db_xref="UniProt/TrEMBL:Q9UIN7"
                     /translation="FKPGTGYSQVSYSSGSGSSLQGASGSSQLGSSSSHSGNSGSHSG
                     SSSSHSSSSSSFQFSSSSFQVGNGSALPTNDNSYRGILNPSQPGQSSSSSQTFGVSSS
                     GQSVSSNQRPCSSDIPDSPCSGGPIVSHSGPYIPSSHSVSGGQRPVVVVVDQHGSGAP
                     GVVQGPPCSNGGLPGKPCPPITSVDKSYGGYEVVGGSSDSYLVPGMTYSKGKIYPVGY
                     FTKENPVKGSPGVPSFAAGPPISEGKYFSSNPIIPSQSAASSAIAFQPVGTGGVQLCG
                     GGSTGSKGPCSPSSSRVPSSSSISSSSGLPYHPCGSASQSPCSPPGTGSFSSSSSSQS
                     SGKIILQPCGSKSSSSGHPCMSVSSLTLTGGPDGSPHPDPSAGAKPCGSSSAGKIPCR
                     SIRDILAQVKPLGPQLADPEVFLPQGELLDSP"
ORIGIN

  Query Match             44.2%;  Score 702;  DB 8;  Length 1327;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1252;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy        315 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCT 374
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          1 TTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCT 60
```

```
Qy        375 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db         61 ACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAACAGCGG 120

Qy        435 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAG 180

Qy        495 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGG 554
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        181 CAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGG 240

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATC 614
              |||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||
Db        241 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAGACCTTTGGGGTATC 300

Qy        615 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTC 360

Qy        675 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTC 734
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        361 TCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATCCCCAGCTCCCACTC 420

Qy        735 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCC 480

Qy        795 TGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCC 540

Qy        855 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAG 600

Qy        915 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCAC 974
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        601 TTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCTGTGGGCTACTTCAC 660

Qy        975 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCAT 720

Qy       1035 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTC 780

Qy       1095 GGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 1154
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCAC 840

Qy       1155 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTC 900

Qy       1215 CAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 1274
              ||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        901 CAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCC 960

Qy       1275 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCT 1334
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        961 ACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGTGGCAAAATCATCCT 1020

Qy       1335 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTT 1080

Qy       1395 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCC 1140

Qy       1455 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCA 1200

Qy       1515 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTT 1260

Qy       1575 ACT 1577
              |||
Db       1261 ACT 1263
```

Search completed: March 22, 2006, 05:12:27
Job time : 8140 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:57:01 ; Search time 6677 Seconds
(without alignments)
11141.450 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1386	87.2	1590	10	AY412935	AY412935 Homo sapi
	2	497	31.3	648	5	BX505094	BX505094 DKFZp686G
	3	497	31.3	673	1	AU134804	AU134804 AU134804
	4	495	31.1	581	3	BP226265	BP226265 BP226265
	5	480	30.2	1590	10	AY412936	AY412936 Pan trogl
	6	351	22.1	531	8	DN995832	DN995832 TC110851
	7	344	21.6	644	7	CV025058	CV025058 2636 Full
	8	305	19.2	685	1	AU297843	AU297843 AU297843
	9	284	17.9	357	2	BF834417	BF834417 CM2-HT096
c	10	245	15.4	541	2	BE140467	BE140467 IL0-HT000
c	11	230	14.5	519	2	BE140469	BE140469 IL0-HT000
	12	213	13.4	742	2	BG676174	BG676174 602622574
	13	204	12.8	354	2	BF834420	BF834420 CM2-HT096
	14	200	12.6	426	1	AU298085	AU298085 AU298085
	15	193	12.1	345	2	BF834425	BF834425 CM2-HT096
c	16	179	11.3	1077	3	BM808976	BM808976 AGENCOURT
c	17	175	11.0	660	1	AI768204	AI768204 wg82d08.x
	18	158	9.9	371	2	BF835362	BF835362 CM2-HT096
	19	80	5.0	500	1	AU297870	AU297870 AU297870
c	20	63	4.0	361	2	BG899994	BG899994 HOA48-1-F
	21	50	3.1	655	6	CF763596	CF763596 CES005522
	22	50	3.1	678	6	CF767753	CF767753 CES006300
	23	49	3.1	173	2	BF834547	BF834547 CM2-HT096
c	24	45	2.8	690	1	AI582314	AI582314 tq66d05.x
	25	40	2.5	121	2	BE140468	BE140468 IL0-HT000
c	26	30	1.9	568	7	CO664466	CO664466 DG40-133e
	27	28	1.8	136	2	BG991881	BG991881 MR2-HT116
	28	27	1.7	544	1	AW786881	AW786881 120502 MA
	29	27	1.7	545	2	BE030827	BE030827 129035 MA
	30	27	1.7	591	2	BG384333	BG384333 303380 MA
	31	26	1.6	367	5	BY772273	BY772273 BY772273
	32	26	1.6	605	1	AI746833	AI746833 ul08a01.y
	33	23	1.4	106	8	T25011	T25011 EST586 Huma
	34	23	1.4	376	1	AI760969	AI760969 wi70g12.x
	35	23	1.4	399	1	AI342924	AI342924 qo36f05.x
	36	23	1.4	404	1	AI421163	AI421163 tf19b11.x
	37	23	1.4	405	1	AI095136	AI095136 qa19d05.x
	38	23	1.4	405	1	AW182910	AW182910 xj65a03.x
	39	23	1.4	409	1	AV625944	AV625944 AV625944
	40	23	1.4	410	2	BE220139	BE220139 hv67f04.x
	41	23	1.4	412	1	AW134534	AW134534 UI-H-BI1-
	42	23	1.4	421	1	AA973512	AA973512 oo46a04.s
	43	23	1.4	458	1	AA993413	AA993413 ot96b08.s
	44	23	1.4	461	1	AV632453	AV632453 AV632453
c	45	23	1.4	505	7	CO981353	CO981353 GM89011A1
c	46	23	1.4	522	3	BQ013461	BQ013461 UI-1-BC1p
	47	23	1.4	526	1	AA460713	AA460713 zx69f03.s
c	48	23	1.4	535	3	BQ012983	BQ012983 UI-1-BC1p
	49	23	1.4	641	6	CF565539	CF565539 1115101B0
	50	23	1.4	643	6	CF559435	CF559435 1115047A0
	51	23	1.4	643	6	CF568566	CF568566 1115126C1
c	52	23	1.4	672	6	CD370084	CD370084 UI-H-FT1-
c	53	23	1.4	690	6	CD743189	CD743189 UI-H-FT1-
	54	23	1.4	703	6	CF566931	CF566931 1115112B1
c	55	23	1.4	720	7	CO108894	CO108894 GR__Eb004
c	56	23	1.4	750	6	CD366695	CD366695 UI-H-FT2-
c	57	22	1.4	190	9	AZ246603	AZ246603 RPCI-23-5
	58	22	1.4	326	1	AW416113	AW416113 51152 MAR
	59	22	1.4	378	1	AA717857	AA717857 vs76f07.r
	60	22	1.4	432	8	H03774	H03774 yj39a06.r1
c	61	22	1.4	484	1	AA700303	AA700303 zj73g09.s
c	62	22	1.4	553	6	CF178850	CF178850 808375 MA
	63	22	1.4	593	5	BX671416	BX671416 BX671416
c	64	22	1.4	606	9	AQ464717	AQ464717 HS_5104_B
	65	22	1.4	609	2	BB615758	BB615758 BB615758
	66	22	1.4	617	7	CF906543	CF906543 A0448E01-
c	67	22	1.4	655	10	AG033139	AG033139 Pan trogl
c	68	22	1.4	696	10	AG319650	AG319650 Mus muscu
c	69	22	1.4	710	10	AG319699	AG319699 Mus muscu
	70	22	1.4	1674	10	AY412937	AY412937 Mus muscu
c	71	21	1.3	107	2	BE342791	BE342791 EST395635
	72	21	1.3	119	8	DR530772	DR530772 WS02733.C
	73	21	1.3	139	8	DR537824	DR537824 WS02753.C
c	74	21	1.3	152	10	CW469462	CW469462 fsbb001f2
	75	21	1.3	166	7	CF906637	CF906637 A0449F08-

ALIGNMENTS

RESULT 1
AY412935
LOCUS AY412935 1590 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens CDSN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY412935
VERSION AY412935.1 GI:39768900
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1590)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1590)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .1590
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1. .>1590
/gene="CDSN"
/locus_tag="HCM4723"
ORIGIN

Query Match 87.2%; Score 1386; DB 10; Length 1590;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1586; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
BX505094
LOCUS BX505094 648 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686G12192_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone DKFZp686G12192 5', mRNA sequence.
ACCESSION BX505094

```
VERSION     BX505094.1  GI:32032777
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
            Wellenreuther,R., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686G12192) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686G12192"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             31.3%;  Score 497;  DB 5;  Length 648;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-237;
  Matches  547;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 88

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 148

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 208

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 268

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 328

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 388

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 448

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 508

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 568

Qy        541 AACTCTTA 548
              ||||||||
Db        569 AACTCTTA 576


RESULT 3
AU134804
LOCUS       AU134804                673 bp    mRNA    linear   EST 02-AUG-2002
DEFINITION  AU134804 PLACE1 Homo sapiens cDNA clone PLACE1000541 5', mRNA
            sequence.
ACCESSION   AU134804
VERSION     AU134804.1  GI:10995343
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .673
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1000541"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match             31.3%;  Score 497;  DB 1;  Length 673;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-237;
  Matches  547;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 88

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 148

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 208

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 268
```

```
Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 328

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 388

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 448

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 508

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 568

Qy        541 AACTCTTA 548
              ||||||||
Db        569 AACTCTTA 576
```

```
RESULT 4
BP226265
LOCUS       BP226265                581 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP226265 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
            clone DMC02127, mRNA sequence.
ACCESSION   BP226265
VERSION     BP226265.1  GI:52099170
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DMC02127"
                     /clone_lib="Sugano cDNA library, dermoid cancer"
                     /note="dermoid cancer"
ORIGIN

  Query Match             31.1%;  Score 495;  DB 3;  Length 581;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-236;
  Matches  545;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCT 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 GGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCT 88

Qy         63 GCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 148

Qy        123 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGA 208

Qy        183 CTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGC 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGC 268

Qy        243 CAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGG 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGG 328

Qy        303 TTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATC 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATC 388

Qy        363 TGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTC 422
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 TGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTC 448

Qy        423 GGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 GGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTT 508

Qy        483 TCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAA 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 TCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAA 568

Qy        543 CTCTTA 548
              ||||||
Db        569 CTCTTA 574
```

```
RESULT 5
AY412936
LOCUS       AY412936                1590 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Pan troglodytes CDSN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412936
VERSION     AY412936.1  GI:39768901
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 1590)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 1590)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .1590
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>1590
                     /gene="CDSN"
                     /locus_tag="HCM4723"
ORIGIN
```

```
  Query Match             30.2%;  Score 480;  DB 10;  Length 1590;
  Best Local Similarity   99.5%;  Pred. No. 2.2e-228;
  Matches  630;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTT 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTT 1017

Qy       1018 GCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 GCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGC 1077

Qy       1078 CAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTC 1137

Qy       1138 TGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCC 1197
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1138 TGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCNAGTCCCC 1197

Qy       1198 AGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCC 1257
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1198 AGCAGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCC 1257

Qy       1258 CAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCG 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 CAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCC 1317

Qy       1318 AGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 AGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377

Qy       1378 ATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 ATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCC 1437

Qy       1438 TCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 TCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATC 1497

Qy       1498 CGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 CGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTC 1557

Qy       1558 CTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||||||
Db       1558 CTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 6
DN995832
LOCUS       DN995832                 531 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC110851 Human breast cancer tissue, large insert, pCMV expression
            library Homo sapiens cDNA clone TC110851 5' similar to Homo sapiens
            corneodesmosin (CDSN), mRNA sequence.
ACCESSION   DN995832
VERSION     DN995832.1  GI:66255659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
            6 Taft Ct. Suite 100
            Rockville, MD 20850
            Tel: (301) 340-3188
            http://www.origene.com
            Seq primer: pCMV6 5prime forward vector primer, OriGene
            Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC110851"
                     /tissue_type="Breast cancer"
                     /clone_lib="Human breast cancer tissue, large insert, pCMV
                     expression library"
                     /note="Organ: Mammary gland (cancer tissue); Vector:
                     pCMV6-XL5; Site_1: EcoR1; Site_2: Xho1/Sal1 compatible end
                     ligatio; Oligo-dT primed reverse transcription optimized
                     for large and GC rich mRNA transcripts, cDNA size
                     selection, optimized ligation for large inserts into
                     mammalian expression vector, random clones selected for
                     end sequence verification of full-length genes"
ORIGIN

  Query Match             22.1%;  Score 351;  DB 8;  Length 531;
  Best Local Similarity   99.6%;  Pred. No. 8.1e-164;
  Matches  451;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          25 ATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCA 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 ATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCA 138

Qy          85 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         139 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 198

Qy         145 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         199 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 258

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         259 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 318

Qy         265 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         319 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 378

Qy         325 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCA 384
               ||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         379 GGAACGGNGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCA 438

Qy         385 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 444
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         439 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 498

Qy         445 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGC 477
               |||||||||||||||||||||||||||||||||
Db         499 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGC 531


RESULT 7
CV025058
LOCUS       CV025058                 644 bp    mRNA    linear   EST 20-AUG-2004
```

```
DEFINITION  2636 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC031993, mRNA sequence.
ACCESSION   CV025058
VERSION     CV025058.1  GI:51482951
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGGCTCGTCTCGGGCA
            BACKWARD: TATGGACTGTCGAGTAACTCTCCTT
            Insert Length: 644   Std Error: 36.00
            Plate: 11068  row: 04  column: B
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 98
            High quality sequence stop: 643
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="mixed"
                     /clone_lib="Full Length cDNA from the Mammalian Gene
                     Collection"
                     /note="Vector: mixed; The ORFs were PCR amplified from the
                     MGC (Mammalian Gene Collection) as of April 2004 and
                     cloned by recombinational Gateway cloning into pDONR223
                     Donor vector. Reference : MGC (Mammalian Gene Collection)
                     Program Team, Generation and Initial Analysis of more than
                     15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                     2002, 99(26), 16899-16903"
ORIGIN

  Query Match             21.6%;  Score 344;  DB 7;  Length 644;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-160;
  Matches  444;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGG 446
              |||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGG 446


RESULT 8
AU297843
LOCUS       AU297843                685 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  AU297843 male adult skin, full-length enriched chimpanzee cDNA
            library Pan troglodytes verus cDNA clone PstA0073 5' similar to
            human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION   AU297843
VERSION     AU297843.1  GI:29532142
KEYWORDS    EST.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N.,
            Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
  TITLE     Analysis of 5'-end sequences of chimpanzee cDNAs
  JOURNAL   Genome Res. 13 (5), 1022-1026 (2003)
   PUBMED   12727913
COMMENT     Contact: Momoki Hirai
            Department of Integrated Biosciences
            The University of Tokyo, Graduate School of Frontier Sciences
            5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
            Tel: 81-4-7136-3688
            Fax: 81-4-7136-3687
            Email: mhirai@k.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Pan troglodytes verus"
                     /mol_type="mRNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone="PstA0073"
                     /sex="male"
                     /tissue_type="skin"
                     /dev_stage="adult"
                     /clone_lib="male adult skin, full-length enriched
                     chimpanzee cDNA library"
ORIGIN

  Query Match             19.2%;  Score 305;  DB 1;  Length 685;
  Best Local Similarity   99.2%;  Pred. No. 8.9e-141;
  Matches  505;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         40 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 126

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        127 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCATGCGTATCACCTCCCCTAACGAC 186

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        187 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 246

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        247 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGCGGTAGCTCCAGC 306

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 366

Qy        340 CAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAG 399
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        367 CAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAG 426

Qy        400 CTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 459
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        427 CTGGGGAGCAGCAGCTCTCACTCAGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCT 486

Qy        460 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGC 519
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGC 546

Qy        520 TCTGCTCTGCCAACCAATGACAACTCTTA 548
              ||||||||||||||||||||||||||||||
Db        547 TCTGCTCTGCCAACCAATGACAACTCTTA 575


RESULT 9
BF834417
LOCUS       BF834417                357 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  CM2-HT0968-111100-500-b11 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF834417
VERSION     BF834417.1  GI:12184792
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-
            111100-500-b11&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 357.
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0968"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             17.9%;  Score 284;  DB 2;  Length 357;
  Best Local Similarity   99.7%;  Pred. No. 2.8e-130;
  Matches  334;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        872 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 931
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 AATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCA 82

Qy        932 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 991
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 TGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGA 142

Qy        992 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 1051
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACT 202

Qy       1052 TCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGC 1111
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        203 TCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCGGCTTCCTCGGCCATTGCGTTCCAGC 262

Qy       1112 CAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCT 1171
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCT 322

Qy       1172 GCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 1206
              |||||||||||||||||||||||||||||||||||
Db        323 GCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 357


RESULT 10
BE140467/c
LOCUS       BE140467                541 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  IL0-HT0007-140599-004 HT0007 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE140467
VERSION     BE140467.1  GI:8603188
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL0-HT0007-140
            599-004&t3=1999-05-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 50
```

```
           High quality sequence stop: 542.
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0007"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             15.4%;  Score 245;  DB 2;  Length 541;
  Best Local Similarity   100.0%;  Pred. No. 9.7e-111;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1346 GCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTG 1405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTG 441

Qy       1406 GGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCA 1465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCA 381

Qy       1466 GCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTC 1525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTC 321

Qy       1526 TGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTC 1585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 TGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTC 261

Qy       1586 CATAA 1590
             |||||
Db        260 CATAA 256


RESULT 11
BE140469/c
LOCUS       BE140469                 519 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  IL0-HT0007-140599-002 HT0007 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE140469
VERSION     BE140469.1  GI:8603190
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br

            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL0-HT0007-140
            599-002&t3=1999-05-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 55
            High quality sequence stop: 134.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0007"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             14.5%;  Score 230;  DB 2;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-103;
  Matches  230;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1361 CTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCT 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCT 460

Qy       1421 CTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGA 1480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 CTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGA 400

Qy       1481 TCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAG 1540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 TCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAG 340

Qy       1541 CTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 CTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 290


RESULT 12
BG676174
LOCUS       BG676174                 742 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602622574F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4747362 5',
            mRNA sequence.
ACCESSION   BG676174
VERSION     BG676174.1  GI:13907570
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10597  row: d  column: 19
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
```

```
    source          1. .742
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4747362"
                    /tissue_type="squamous cell carcinoma"
                    /lab_host="DH10B (T1 phage-resistant)"
                    /clone_lib="NCI_CGAP_Skn4"
                    /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                    Average insert size 1.5kb.  Library  constructed by Life
                    Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             13.4%;  Score 213;  DB 2;  Length 742;
  Best Local Similarity   100.0%;  Pred. No. 1e-94;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          86 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 145
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 151

Qy         146 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         152 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 211

Qy         206 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 265
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         212 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 271

Qy         266 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGG 298
               ||||||||||||||||||||||||||||||||
Db         272 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGG 304


RESULT 13
BF834420
LOCUS       BF834420                354 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  CM2-HT0968-111100-500-d11 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF834420
VERSION     BF834420.1  GI:12184797
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-
            111100-500-d11&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 25
            High quality sequence stop: 354.
FEATURES            Location/Qualifiers
     source         1. .354
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /dev_stage="Adult"
                    /clone_lib="HT0968"
                    /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                    Site_2: SmaI; A mini-library was made by cloning products
                    derived from  ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 -  Ludwig Institute for Cancer Research)
                    profiles into  the pUC 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."
ORIGIN

  Query Match             12.8%;  Score 204;  DB 2;  Length 354;
  Best Local Similarity   99.3%;  Pred. No. 3.2e-90;
  Matches  304;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          49 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 108

Qy         961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         109 GTGGGCGACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 168

Qy        1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         169 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 228

Qy        1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
               |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 TCGGCAGCGTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 288

Qy        1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         289 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 348

Qy        1201 AGTTCT 1206
               ||||||
Db         349 AGTTCT 354


RESULT 14
AU298085
LOCUS       AU298085                426 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  AU298085 male adult skin, full-length enriched chimpanzee cDNA
            library Pan troglodytes verus cDNA clone PstA2015 5' similar to
            human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION   AU298085
VERSION     AU298085.1  GI:29532384
KEYWORDS    EST.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N.,
            Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
  TITLE     Analysis of 5'-end sequences of chimpanzee cDNAs
  JOURNAL   Genome Res. 13 (5), 1022-1026 (2003)
   PUBMED   12727913
COMMENT     Contact: Momoki Hirai
            Department of Integrated Biosciences
            The University of Tokyo, Graduate School of Frontier Sciences
            5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
            Tel: 81-4-7136-3688
            Fax: 81-4-7136-3687
            Email: mhirai@k.u-tokyo.ac.jp.
FEATURES            Location/Qualifiers
```

```
     source          1. .426
                     /organism="Pan troglodytes verus"
                     /mol_type="mRNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone="PstA2015"
                     /sex="male"
                     /tissue_type="skin"
                     /dev_stage="adult"
                     /clone_lib="male adult skin, full-length enriched
                     chimpanzee cDNA library"
ORIGIN

  Query Match             12.6%;  Score 200;  DB 1;  Length 426;
  Best Local Similarity   99.3%;  Pred. No. 3.3e-88;
  Matches  300;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         40 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GGGCACGGGATGATGGCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAG 126

Qy        100 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGAC 159
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        127 AGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCATGCGTATCACCTCCCCTAACGAC 186

Qy        160 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 CCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGT 246

Qy        220 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        247 TCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGCGGTAGCTCCAGC 306

Qy        280 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCC 366

Qy        340 CA 341
              ||
Db        367 CA 368


RESULT 15
BF834425
LOCUS       BF834425                345 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  CM2-HT0968-111100-500-g03 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF834425
VERSION     BF834425.1  GI:12184811
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-
            111100-500-g03&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 345.
FEATURES    Location/Qualifiers
     source 1. .345
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="HT0968"
            /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from  ORESTES PCR (U.S. Letters Patent application
            No. 196,716 -  Ludwig Institute for Cancer Research)
            profiles into  the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
ORIGIN

  Query Match             12.1%;  Score 193;  DB 2;  Length 345;
  Best Local Similarity   99.6%;  Pred. No. 1e-84;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        963 GGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGC 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGC 161

Qy       1023 TGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTC 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        162 TGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCGGTC 221

Qy       1083 GGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGG 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 GGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGG 281

Qy       1143 AGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAG 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 AGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAG 341

Qy       1203 TTCT 1206
              ||||
Db        342 TTCT 345


RESULT 16
BM808976/c
LOCUS       BM808976                1077 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  AGENCOURT_6582710 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471455
            5', mRNA sequence.
ACCESSION   BM808976
VERSION     BM808976.1  GI:19125799
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1077)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
            http://image.llnl.gov
            Plate: LLCM1979  row: o  column: 08
            High quality sequence start: 41
            High quality sequence stop: 618.
FEATURES              Location/Qualifiers
     source           1. .1077
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5471455"
                      /tissue_type="amelanotic melanoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_41"
                      /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Library constructed by Ling Hong in the
                      laboratory of Gerald M. Rubin (University of California,
                      Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                      Superscript II RT (Life Technologies). Note: this is a
                      NIH_MGC Library."
ORIGIN

  Query Match             11.3%;  Score 179;  DB 3;  Length 1077;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-77;
  Matches  229;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         275 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 216

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         215 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 156

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         155 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGAT 96

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 46


RESULT 17
AI768204/c
LOCUS       AI768204                660 bp    mRNA    linear   EST 20-DEC-1999
DEFINITION  wg82d08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2371599 3' similar to SW:CDSN_HUMAN Q15517 CORNEODESMOSIN
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI768204
VERSION     AI768204.1  GI:5234713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 660)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1598   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 396.
FEATURES              Location/Qualifiers
     source           1. .660
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2371599"
                      /lab_host="DH10B"
                      /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                      /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                      a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                      Equal amounts of plasmid DNA from five normalized
                      libraries were mixed, and ss circles were made in vitro.
                      Following HAP purification, this DNA was used as tracer in
                      a subtractive hybridization reaction. The driver was
                      PCR-amplified cDNAs from pools of 5,000 clones made from
                      the same 5 libraries. The pools consisted of the following
                      libraries and cloneIDs: Soares NbHSF pool 1:
                      309384-310919, 323208-325895 Soares Nb2HP pool 1:
                      145032-147335, 147720-148103, 148872-149255, 15002 -
                      150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                      758280-760583, 772104-774407 Soares NbHPA pool 1:
                      304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                      pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                      Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match             11.0%;  Score 175;  DB 1;  Length 660;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-75;
  Matches  175;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1404 TGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTC 1463
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         646 TGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTC 587

Qy        1464 CAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCC 1523
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         586 CAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCC 527

Qy        1524 TCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTC 1578
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         526 TCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTC 472


RESULT 18
BF835362
LOCUS       BF835362                371 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  CM2-HT0968-141100-504-e05 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF835362
VERSION     BF835362.1  GI:12186401
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 371)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
```

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-141100-504-e05&t3=2000-11-14&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 63
High quality sequence stop: 371.

```
FEATURES             Location/Qualifiers
     source          1. .371
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0968"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             9.9%;  Score 158;  DB 2;  Length 371;
  Best Local Similarity   99.5%;  Pred. No. 3.5e-67;
  Matches  208;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        950 AAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCC 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 AAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCC 174

Qy       1010 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCA 234

Qy       1070 TCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 1129
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TCCCCAGCCGGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGG 294

Qy       1130 TCCAGCTCTGTGGAGGCGGCTCCACGGGC 1158
              ||||||||||||||||||||||||||||
Db        295 TCCAGCTCTGTGGAGGCGGCTCCACGGGC 323


RESULT 19
AU297870
LOCUS        AU297870                500 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION   AU297870 male adult skin, full-length enriched chimpanzee cDNA
             library Pan troglodytes verus cDNA clone PstA0195 5' similar to
             human RefSeq mRNA NM_001264, mRNA sequence.
ACCESSION    AU297870
VERSION      AU297870.1  GI:29532169
KEYWORDS     EST.
SOURCE       Pan troglodytes verus
  ORGANISM   Pan troglodytes verus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Pan.
REFERENCE    1  (bases 1 to 500)
  AUTHORS    Sakate,R., Osada,N., Hida,M., Sugano,S., Hayasaka,I., Shimohira,N.,
             Yanagi,S., Suto,Y., Hashimoto,K. and Hirai,M.
  TITLE      Analysis of 5'-end sequences of chimpanzee cDNAs
  JOURNAL    Genome Res. 13 (5), 1022-1026 (2003)
   PUBMED    12727913
COMMENT      Contact: Momoki Hirai
             Department of Integrated Biosciences
             The University of Tokyo, Graduate School of Frontier Sciences
             5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
             Tel: 81-4-7136-3688
             Fax: 81-4-7136-3687
             Email: mhirai@k.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .500
                     /organism="Pan troglodytes verus"
                     /mol_type="mRNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone="PstA0195"
                     /sex="male"
                     /tissue_type="skin"
                     /dev_stage="adult"
                     /clone_lib="male adult skin, full-length enriched
                     chimpanzee cDNA library"
ORIGIN

  Query Match             5.0%;  Score 80;  DB 1;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 4e-28;
  Matches   80;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        138 GCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 GCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAG 224

Qy        198 TAGCTACAGTGGCTCCAGCA 217
              ||||||||||||||||||||
Db        225 TAGCTACAGTGGCTCCAGCA 244


RESULT 20
BG899994/c
LOCUS        BG899994                361 bp    mRNA    linear   EST 06-NOV-2001
DEFINITION   HOA48-1-F7.R HOA (Human Osteoarthritic Cartilage) Homo sapiens
             cDNA, mRNA sequence.
ACCESSION    BG899994
VERSION      BG899994.1  GI:14310243
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE    1  (bases 1 to 361)
  AUTHORS    Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
             Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
             Lark,M.W.
  TITLE      Identification and initial characterization of 5000 expressed
             sequenced tags (ESTs) each from adult human normal and
             osteoarthritic cartilage cDNA libraries
  JOURNAL    Osteoarthr. Cartil. 9 (7), 641-653 (2001)
   PUBMED    11597177
COMMENT      Contact: Sanjay Kumar
             UW2109
             GlaxoSmithKline
             709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
             Tel: 610-270-7245
             Fax: 610-270-5598
             Email: sanjay_kumar-1@gsk.com
             Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .361
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="cartilage"
                     /lab_host="E.coli DH10 B"
                     /clone_lib="HOA (Human Osteoarthritic Cartilage)"
                     /note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
                     Directional"
ORIGIN

  Query Match             4.0%;  Score 63;  DB 2;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-19;
  Matches   63;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        465 GAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGC 154
```

```
Qy        525 TCT 527
              |||
Db        153 TCT 151


RESULT 21
CF763596
LOCUS       CF763596                 655 bp    mRNA    linear   EST 17-OCT-2003
DEFINITION  CES005522 Bos taurus skin cDNA library Bos taurus cDNA clone
            CCL005522 5', mRNA sequence.
ACCESSION   CF763596
VERSION     CF763596.1  GI:37712814
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Wang,Y.H., McWilliam,S. and Lehnert,S.
  TITLE     Transcription profiling of cattle skin
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dr Yonghong Wang
            Functional Genomics Lab
            CSIRO Livestock Industries
            Level 5, Queensland Biosciences Precinct, University of Queensland,
            306 Carmody Road St.Lucia QLD Australia
            Tel: 07 3214 2445
            Fax: 07 3214 2685
            Email: Yonghong.Wang@csiro.au
            Plate: 65  row: D  column: 04.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Hereford Shorthorn"
                     /db_xref="taxon:9913"
                     /clone="CCL005522"
                     /sex="female"
                     /tissue_type="pooled"
                     /dev_stage="Adult"
                     /lab_host="XL1-BlueMRF'strain"
                     /clone_lib="Bos taurus skin cDNA library"
                     /note="Organ: skin; Vector: Uni-ZAPXR; Site_1: EcoRI;
                     Site_2: Xho I; Library made from pooled skin of adult
                     female Hereford-Shorthorn."
ORIGIN

  Query Match             3.1%;  Score 50;  DB 6;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1516 GTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCA 161


RESULT 22
CF767753
LOCUS       CF767753                 678 bp    mRNA    linear   EST 17-OCT-2003
DEFINITION  CES006300 Bos taurus skin cDNA library Bos taurus cDNA clone
            CCL006300 5', mRNA sequence.
ACCESSION   CF767753
VERSION     CF767753.1  GI:37716972
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Wang,Y.H., McWilliam,S. and Lehnert,S.
  TITLE     Transcription profiling of cattle skin
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dr Yonghong Wang
            Functional Genomics Lab
            CSIRO Livestock Industries
            Level 5, Queensland Biosciences Precinct, University of Queensland,
            306 Carmody Road St.Lucia QLD Australia
            Tel: 07 3214 2445
            Fax: 07 3214 2685
            Email: Yonghong.Wang@csiro.au
            Plate: 72  row: F  column: 03.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Hereford Shorthorn"
                     /db_xref="taxon:9913"
                     /clone="CCL006300"
                     /sex="female"
                     /tissue_type="pooled"
                     /dev_stage="Adult"
                     /lab_host="XL1-BlueMRF'strain"
                     /clone_lib="Bos taurus skin cDNA library"
                     /note="Organ: skin; Vector: Uni-ZAPXR; Site_1: EcoRI;
                     Site_2: Xho I; Library made from pooled skin of adult
                     female Hereford-Shorthorn."
ORIGIN

  Query Match             3.1%;  Score 50;  DB 6;  Length 678;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-13;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1516 GTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCA 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 GTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCA 582


RESULT 23
BF834547
LOCUS       BF834547                 173 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  CM2-HT0968-111100-512-b03 HT0968 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF834547
VERSION     BF834547.1  GI:12185204
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 173)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM2&t2=CM2-HT0968-
            111100-512-b03&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 14
```

```
          High quality sequence stop: 170.
FEATURES             Location/Qualifiers
     source          1. .173
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0968"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             3.1%;  Score 49;  DB 2;  Length 173;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-12;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        990 GAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCT 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCT 112


RESULT 24
AI582314/c
LOCUS       AI582314                690 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION  tq66d05.x1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2213769 3'
            similar to TR:Q15517 Q15517 S PROTEIN. ;, mRNA sequence.
ACCESSION   AI582314
VERSION     AI582314.1  GI:4568211
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 690)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 777    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2213769"
                     /tissue_type="squamous cell carcinoma, poorly
                     differentiated (4 pooled tumors, including primary and
                     metastatic)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Lu19"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     pooled lung tumor tissue, and was then primed with a Not I
                     - oligo(dT) primer. Double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I  and Eco RI sites of the modified
                     pT7T3 vector.  Library went througb one round of
                     normalization. Library constructed by Bento Soares and M.
                     Fatima Bonaldo. "
ORIGIN

  Query Match             2.8%;  Score 45;  DB 1;  Length 690;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1487 GCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGC 1531
              |||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGC 498


RESULT 25
BE140468
LOCUS       BE140468                121 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  IL0-HT0007-140599-003 HT0007 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE140468
VERSION     BE140468.1  GI:8603189
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 121)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL0-HT0007-140
            599-003&t3=1999-05-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 41
            High quality sequence stop: 122.
FEATURES             Location/Qualifiers
     source          1. .121
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0007"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             2.5%;  Score 40;  DB 2;  Length 121;
  Best Local Similarity   100.0%;  Pred. No. 4e-08;
```

```
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1311 CCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
              ||||||||||||||||||||||||||||||||||||||||
Db         15 CCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 54


RESULT 26
CO664466/c
LOCUS       CO664466                568 bp    mRNA    linear   EST 23-JUL-2004
DEFINITION  DG40-133e8 DG40-uterus Canis familiaris cDNA 3', mRNA sequence.
ACCESSION   CO664466
VERSION     CO664466.1  GI:50581256
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Schlueter,T., Hermanns,J., Weindel,M., Schuette,D., Kranz,H.,
            Henrich,J. and Loebbert,R.
  TITLE     Dog arrayTAG cDNA clone collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Thomas Schlueter
            LION bioscience AG
            Waldhoferstrasse 98, D-69123 Heidelberg, Germany
            Tel: +49 6221 4038 150
            Fax: +49 6221 4038 290
            Email: Thomas.Schlueter@lionbioscience.com.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="Beagle"
                     /db_xref="taxon:9615"
                     /tissue_type="uterus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="DG40-uterus"
                     /note="Organ: uterus; Vector: Dog pBluescript LION"
ORIGIN

  Query Match             1.9%;  Score 30;  DB 7;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 0.0044;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1512 CCAAGTGAAGCCTCTGGGGCCCCAGCTAGC 1541
              ||||||||||||||||||||||||||||||
Db        529 CCAAGTGAAGCCTCTGGGGCCCCAGCTAGC 500


RESULT 27
BG991881
LOCUS       BG991881                136 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  MR2-HT1160-310101-006-b06 HT1160 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG991881
VERSION     BG991881.1  GI:14395951
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 136)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR2&t2=MR2-HT1160-
            310101-006-b06&t3=2001-01-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 29
            High quality sequence stop: 129.
FEATURES             Location/Qualifiers
     source          1. .136
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT1160"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             1.8%;  Score 28;  DB 2;  Length 136;
  Best Local Similarity   100.0%;  Pred. No. 0.041;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1039 GAGGGCAAATACTTCTCCAGCAACCCCA 1066
              ||||||||||||||||||||||||||||
Db         84 GAGGGCAAATACTTCTCCAGCAACCCCA 111


RESULT 28
AW786881
LOCUS       AW786881                544 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  120502 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW786881
VERSION     AW786881.1  GI:7843657
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
            Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
            EST cluster assembly
  JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
  PUBMED    12226715
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
```

```
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 39  row: J  column: 23
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .544
                      /organism="Sus scrofa"
                      /mol_type="mRNA"
                      /db_xref="taxon:9823"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="MARC 1PIG"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from day 11, 13, 15, 20,
                      and 30 embryos."
ORIGIN

  Query Match             1.7%;  Score 27;  DB 1;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        344 TCAGCTACTCCTCCGGATCTGGCTCTA 370
              |||||||||||||||||||||||||||
Db        411 TCAGCTACTCCTCCGGATCTGGCTCTA 437


RESULT 29
BE030827
LOCUS       BE030827                545 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  129035 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BE030827
VERSION     BE030827.1  GI:8325836
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
            Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
            EST cluster assembly
  JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
   PUBMED   12226715
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 60  row: J  column: 2
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .545
                      /organism="Sus scrofa"
                      /mol_type="mRNA"
                      /db_xref="taxon:9823"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="MARC 1PIG"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from day 11, 13, 15, 20,
                      and 30 embryos."
ORIGIN

  Query Match             1.7%;  Score 27;  DB 2;  Length 545;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        344 TCAGCTACTCCTCCGGATCTGGCTCTA 370
              |||||||||||||||||||||||||||
Db         76 TCAGCTACTCCTCCGGATCTGGCTCTA 102


RESULT 30
BG384333
LOCUS       BG384333                591 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  303380 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BG384333
VERSION     BG384333.1  GI:13308805
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
            Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
            EST cluster assembly
  JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
   PUBMED   12226715
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 91  row: B  column: 3
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .591
                      /organism="Sus scrofa"
                      /mol_type="mRNA"
                      /db_xref="taxon:9823"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="MARC 1PIG"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from day 11, 13, 15, 20,
                      and 30 embryos."
ORIGIN

  Query Match             1.7%;  Score 27;  DB 2;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        344 TCAGCTACTCCTCCGGATCTGGCTCTA 370
              |||||||||||||||||||||||||||
Db        420 TCAGCTACTCCTCCGGATCTGGCTCTA 446


RESULT 31
BY772273
LOCUS       BY772273                367 bp    mRNA    linear   EST 23-MAR-2004
DEFINITION  BY772273 RIKEN full-length enriched, 17.5 days embryo whole body
            Mus musculus cDNA clone L930062A18 5', mRNA sequence.
ACCESSION   BY772273
VERSION     BY772273.1  GI:39698911
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
            Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
            Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
            Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
            Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
            Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
            Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
            Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
            Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
  TITLE     Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
  JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
   PUBMED   12819125
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="L930062A18"
                     /tissue_type="whole body"
                     /dev_stage="17.5 days embryo"
                     /clone_lib="RIKEN full-length enriched, 17.5 days embryo
                     whole body"
ORIGIN

  Query Match             1.6%;  Score 26;  DB 5;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 0.43;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        113 TCTCAGACCCCTGTAAGGACCCCACG 138
              ||||||||||||||||||||||||||
Db        143 TCTCAGACCCCTGTAAGGACCCCACG 168


RESULT 32
AI746833
LOCUS       AI746833                605 bp    mRNA    linear   EST 22-JUN-1999
DEFINITION  ul08a01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
            IMAGE:2076168 5' similar to SW:CDSN_HUMAN Q15517 CORNEODESMOSIN
            PRECURSOR ;, mRNA sequence.
ACCESSION   AI746833
VERSION     AI746833.1  GI:5125097
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
            Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
            Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:994604
            Seq primer: custom primer used
            High quality sequence stop: 519.
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2076168"
                     /dev_stage="embryo, 14 dpc"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse embryo mewa"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.  Size selection was performed to exclude fragments
                     <1.5kb.  Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science).
                     Custom primers  for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA."
ORIGIN

  Query Match             1.6%;  Score 26;  DB 1;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        113 TCTCAGACCCCTGTAAGGACCCCACG 138
              ||||||||||||||||||||||||||
Db        136 TCTCAGACCCCTGTAAGGACCCCACG 161


RESULT 33
T25011
LOCUS       T25011                  106 bp    mRNA    linear   EST 24-AUG-1995
DEFINITION  EST586 Human colorectal cancer Homo sapiens cDNA clone 21C2, mRNA
            sequence.
ACCESSION   T25011
VERSION     T25011.1  GI:534636
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Frigerio,J.-M., Berthezene,P., Garrido,P., Ortiz,E., Barthellemy,S,
            Vasseur,S., Sastre,B., Seleznieff,I., Dagorn,J.-C. and
            Iovanna,J.-L.
  TITLE     Analysis of 2166 clones from a human colorectal cancer cDNA library
            by partial sequencing
  JOURNAL   Hum. Mol. Genet. 4, 37-43 (1995)
   PUBMED   7711732
COMMENT     Contact: Iovanna JL
            U.315 INSERM
            46 Bd de la Gaye, F-13009 Marseille, France.
```

```
            Tel: (33) 91 82 03 15
            Fax: (33) 91 26 62 19
            Email: dagorn@arthur.citi2.fr
            This sequence is one of a series obtained by systematic  sequencing
            of a colorectal cancer cDNA library.
            Seq primer: M13 Forward.
FEATURES                Location/Qualifiers
     source             1. .106
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="21C2"
                        /lab_host="E. coli NM522"
                        /clone_lib="Human colorectal cancer"
                        /note="Vector: pT7T3D; Site_1: Eco RI; Site_2: Not I; mRNA
                        was purified from a colorectal tumour of an adult male.
                        cDNA was constructed and cloned into the pT7T3D phagemid
                        following the manufacturer intructions (Pharmacia)."
ORIGIN

  Query Match             1.4%;  Score 23;  DB 8;  Length 106;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db         18 GAGGCCTGTGGTGGTGGTGGTGG 40


RESULT 34
AI760969
LOCUS       AI760969                376 bp    mRNA    linear   EST 21-DEC-1999
DEFINITION  wi70g12.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398726 3'
            similar to contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AI760969
VERSION     AI760969.1  GI:5176636
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 872   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 371.
FEATURES                Location/Qualifiers
     source             1. .376
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2398726"
                        /tissue_type="2 pooled tumors (clear cell type)"
                        /lab_host="DH10B"
                        /clone_lib="NCI_CGAP_Kid12"
                        /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                        Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                        prepared, and ss circles were made in vitro. Following HAP
                        purification, this DNA was used as tracer in a subtractive
                        hybridization reaction. The driver was PCR-amplified cDNAs
                        from a pool of 5,000 clones made from the same library
                        (cloneIDs 1323912-1325831, 1471368-1472903 and
                        1492104-1493255). Subtraction by Bento Soares and M.
                        Fatima Bonaldo.  "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        341 GAGGCCTGTGGTGGTGGTGGTGG 363


RESULT 35
AI342924
LOCUS       AI342924                399 bp    mRNA    linear   EST 01-FEB-1999
DEFINITION  qo36f05.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1910625 3',
            mRNA sequence.
ACCESSION   AI342924
VERSION     AI342924.1  GI:4080130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 896   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 391.
FEATURES                Location/Qualifiers
     source             1. .399
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:1910625"
                        /tissue_type="carcinoid"
                        /lab_host="DH10B"
                        /clone_lib="NCI_CGAP_Lu5"
                        /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                        modified polylinker; 1st strand cDNA was prepared from
                        neuroendocrine lung carcinoid, and was then primed with a
                        Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                        to Eco RI adaptors (Pharmacia), digested with Not I and
                        cloned into the Not I and Eco RI sites of the modified
                        pT7T3 vector.  Library is normalized. Library was
                        constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 399;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
```

```
                    ||||||||||||||||||||||
Db        372 GAGGCCTGTGGTGGTGGTGGTGG 394


RESULT 36
AI421163
LOCUS       AI421163                404 bp    mRNA    linear   EST 28-MAR-1999
DEFINITION  tf19b11.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2096637 3',
            mRNA sequence.
ACCESSION   AI421163
VERSION     AI421163.1  GI:4267094
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 908   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 391.
FEATURES             Location/Qualifiers
     source          1. .404
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2096637"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn23"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 404;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        379 GAGGCCTGTGGTGGTGGTGGTGG 401


RESULT 37
AI095136
LOCUS       AI095136                405 bp    mRNA    linear   EST 18-AUG-1998
DEFINITION  qa19d05.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:1687209 3',
            mRNA sequence.
ACCESSION   AI095136
VERSION     AI095136.1  GI:3434112
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 403.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1687209"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn23"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        379 GAGGCCTGTGGTGGTGGTGGTGG 401


RESULT 38
AW182910
LOCUS       AW182910                405 bp    mRNA    linear   EST 18-NOV-1999
DEFINITION  xj65a03.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2662060 3', mRNA sequence.
ACCESSION   AW182910
VERSION     AW182910.1  GI:6451370
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 394.
FEATURES              Location/Qualifiers
     source           1. .405
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2662060"
                      /lab_host="DH10B"
                      /clone_lib="Soares_NFL_T_GBC_S1"
                      /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                      a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                      Equal amounts of plasmid DNA from three normalized
                      libraries (fetal lung NbHL19W, testis NHT, and B-cell
                      NCI_CGAP_GCB1) were mixed, and ss circles were made in
                      vitro. Following HAP purification, this DNA was used as
                      tracer in a subtractive hybridization reaction. The driver
                      was PCR-amplified cDNAs from pools of 5,000 clones made
                      from the same 3 libraries. The pools consisted of
                      I.M.A.G.E. clones 297480-302087, 682632-687239,
                      726408-728711, and 729096-731399. Subtraction by Bento
                      Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        379 GAGGCCTGTGGTGGTGGTGGTGG 401


RESULT 39
AV625944
LOCUS       AV625944                409 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV625944 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
            reinhardtii cDNA clone LC100b04_r 5', mRNA sequence.
ACCESSION   AV625944
VERSION     AV625944.1  GI:10775121
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
            Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 409)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
   PUBMED   11089912
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES              Location/Qualifiers
     source           1. .409
                      /organism="Chlamydomonas reinhardtii"
                      /mol_type="mRNA"
                      /strain="C9"
                      /db_xref="taxon:3055"
                      /clone="LC100b04_r"
                      /clone_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"
                      /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                      XhoI; The cDNA library was constructed from cells cultured
                      in a carbon stress acclimatized condition in which carbon
                      dioxide concentration in the bubbling gas was changed from
                      5% to 0.04%"
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 409;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        248 GCTCTGGTGGTGGCTCCAGTGGT 270
              |||||||||||||||||||||||
Db        354 GCTCTGGTGGTGGCTCCAGTGGT 376


RESULT 40
BE220139
LOCUS       BE220139                410 bp    mRNA    linear   EST 03-JUL-2000
DEFINITION  hv67f04.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178495 3'
            similar to contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   BE220139
VERSION     BE220139.1  GI:8907457
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 404.
FEATURES              Location/Qualifiers
     source           1. .410
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3178495"
                      /tissue_type="carcinoid"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Lu24"
                      /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Plasmid DNA from the normalized
                      library NCI_CGAP_Lu5 was prepared, and ss circles were
                      made in vitro. Following HAP purification, this DNA was
                      used as tracer in a subtractive hybridization reaction.
                      The driver was PCR-amplified cDNAs from a pool of 5,000
                      clones made from the same library (cloneIDs
                      1414920-1417991 and 1520904-1522439). Subtraction by Bento
                      Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        379 GAGGCCTGTGGTGGTGGTGGTGG 401
```

```
RESULT 41
AW134534
LOCUS       AW134534                412 bp    mRNA    linear   EST 28-OCT-1999
DEFINITION  UI-H-BI1-abv-g-05-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
            IMAGE:2713161 3', mRNA sequence.
ACCESSION   AW134534
VERSION     AW134534.1  GI:6138087
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NCI-CGAP clone distribution
            information can be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html The following repetitive
            elements were found in this cDNA sequence: 385-412,
            >(TGG)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2713161"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub3"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub3 library is a subtracted library derived from
                     the  NCI_CGAP_Sub1 library, which is a subtracted library
                     derived from BI.  BI constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP  libraries:
                     NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
                     NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
                     NCI_CGAP_Kid3,  NCI_CGAP_Kid11, NCI_CGAP_Lym2,
                     NCI_CGAP_Br2, NCI_CGAP_Co8,  NCI_CGAP_CLL1, NCI_CGAP_Lei2,
                     NCI_CGAP_Brn23, NCI_CGAP_Lu5,  NCI_CGAP_Lu24,
                     NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
                     NCI_CGAP_Brn25.  These 21 libraries were pooled and a
                     single-stranded DNA  preparation of the resulting mixture
                     was used as a tracer in a subtractive  hybridization with
                     a driver whose composition is detailed below:
                     NCI_CGAP_Kid3 pool 1     LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1
                     LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs
                     1323912-1325831, 1471368-1472903, 1492104-1493255);
                     NCI_CGAP_Lu5 pool 1     LLAM 3575-3582, 3851-3854 (IMAGE
                     CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
                     pool 1     LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
                     CloneIDs 1257096-1258631, 1469064-1470983,
                     1475592-1476743); NCI_CGAP_Pr22 pool 1     LLAM 2457-2459,
                     2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
                     1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
                     LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
                     1144584-1145351). Subtraction was performed as previously
                     described [Bonaldo, Lennon &  Soares (1996): Normalization
                     and Subtraction: Two Approaches To  Facilitate Gene
                     Discovery. Genome Research 6, 791-806.
                     TAG_SEQ=None found"
ORIGIN
```

```
  Query Match             1.4%;  Score 23;  DB 1;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        378 GAGGCCTGTGGTGGTGGTGGTGG 400


RESULT 42
AA973512
LOCUS       AA973512                421 bp    mRNA    linear   EST 17-JUN-1998
DEFINITION  oo46a04.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1569198 3'
            similar to contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AA973512
VERSION     AA973512.1  GI:3148692
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 945   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1569198"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu5"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     neuroendocrine lung carcinoid, and was then primed with a
                     Not I - oligo(dT) primer. Double-stranded cDNA was ligated
                     to Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector.  Library is normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
```

```
  Query Match             1.4%;  Score 23;  DB 1;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        376 GAGGCCTGTGGTGGTGGTGGTGG 398


RESULT 43
AA993413
LOCUS       AA993413                458 bp    mRNA    linear   EST 27-AUG-1998
DEFINITION  ot96b08.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
```

```
            IMAGE:1624599 3' similar to contains element MSR1 repetitive
            element ;, mRNA sequence.
ACCESSION   AA993413
VERSION     AA993413.1  GI:3179958
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 809   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 409.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1624599"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        377 GAGGCCTGTGGTGGTGGTGGTGG 399


RESULT 44
AV632453
LOCUS       AV632453                461 bp    mRNA    linear   EST 15-DEC-2000
DEFINITION  AV632453 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC008d09_r 5', mRNA sequence.
ACCESSION   AV632453
VERSION     AV632453.1  GI:10775773
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
            Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
  TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
  JOURNAL   DNA Res. 7 (5), 305-307 (2000)
   PUBMED   11089912
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="C9"
                     /db_xref="taxon:3055"
                     /clone="HC008d09_r"
                     /clone_lib="Chlamydomonas reinhardtii 5% CO2"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from cells cultured
                     in a medium with bubbling air containing 5% carbon
                     dioxide"
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 461;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        248 GCTCTGGTGGTGGCTCCAGTGGT 270
              |||||||||||||||||||||||
Db        401 GCTCTGGTGGTGGCTCCAGTGGT 423


RESULT 45
CO981353/c
LOCUS       CO981353                505 bp    mRNA    linear   EST 13-SEP-2004
DEFINITION  GM89011A1D03.r1 Gm-r1089 Glycine max cDNA clone Gm-r1089-3989 3,
            mRNA sequence.
ACCESSION   CO981353
VERSION     CO981353.1  GI:51337487
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 505)
  AUTHORS   Vodkin,L., Shoemaker,R., Keim,P., Polacco,J.C., Retzel,E.,
            Khanna,A., Shealy,R., Clough,S., Thibaud-Nissen,F., Coryell,V.,
            Erpelding,J., Gonzalez,D.O., Stromvik,M., Rodriguez-Huete,A.M.,
            Schweitzer,P., Gong,G. and Liu,L.
  TITLE     A Functional Genomics Program for Soybean (NSF 9872565) (2004)
  JOURNAL   Unpublished (2004)
COMMENT     Other_ESTs: BG363687 corresponding to Gm-c1051-2379 (5')
            Contact: Vodkin, L.O., PI, A Functional Genomics Program for
            Soybean (NSF 9872565)
            Lewin, H. A., Director, Keck Center for Comparative and Functional
            Genomics
            University of Illinois
            Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
            Tel: (217) 244-6147
            Fax: (217) 333-4582
            Email: l-vodkin@uiuc.edu
            Plate: GM89011A1  row: D  column: 03
            Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'
            High quality sequence stop: 505.
FEATURES             Location/Qualifiers
     source          1. .505
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="Gm-r1089-3989"
                     /clone_lib="Gm-r1089"
                     /note="The library Gm-r1089 is a sequence-driven, reracked
                     set of 9,216 low redundancy clones selected from 38
                     different cDNA libraries constructed from various tissues
                     and stages of development of soybean including 973 cDNAs
                     from germinating cotyledons(source library Gm-c1069,
                     Gm-c1076, and Gm-c1077); 1,465 cDNAs from various tissue
```

and organ systems of the adult plant; 476 cDNAs from adult stem tissue (source library Gm-c1062); 1340 cDNAs from tissue culture derived somatic embryos (source libraries Gm-c1036 and Gm-c1075); 2918 cDNAs from hypocotyls or young seedlings; 742 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses(source libraries Gm-c1065, Gm-c1066, and Gm-c1068); 839 cDNAs from young leaves or hypocotyls exposed to bacterial and fungal pathogens (source libraries Gm-c1072, Gm-c1073, Gm-c1074; and Gm-c1084); and 463 from roots of young plants grown in hydroponic media without phosphate (source library Gm-c1087). The 5' ESTs of the source clones from the different libraries were used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1089 and the cDNA clones of the reracked Gm-r1089 library were then sequenced at the 3' end. The unigene selection and 3' sequencing was funded by NSF Plant Genome project #9872565 (http://soybeangenomics.cropsci.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project, http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/. Reracking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/keck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1089 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

```
  Query Match             1.4%;  Score 23;  DB 7;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        114 GAGGCCTGTGGTGGTGGTGGTGG 92
```

RESULT 46
BQ013461/c

LOCUS	BQ013461 522 bp mRNA linear EST 26-MAR-2002
DEFINITION	UI-1-BC1p-ayx-g-08-0-UI.s1 NCI_CGAP_Pl3 Homo sapiens cDNA clone UI-1-BC1p-ayx-g-08-0-UI 3', mRNA sequence.
ACCESSION	BQ013461
VERSION	BQ013461.1 GI:19738362
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 522)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Steven Brown cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu The following repetitive elements were found in this cDNA sequence: 410-436, >(TGG)n#Simple_repeat (matched compliment) Seq primer: M13 FORWARD POLYA=Yes.

FEATURES Location/Qualifiers

```
     source          1. .522
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BC1p-ayx-g-08-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Pl3"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl3 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human 8 week
                     TAG_LIB=UI-1-BC1p
                     TAG_SEQ=GA"
ORIGIN
```

```
  Query Match             1.4%;  Score 23;  DB 3;  Length 522;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        443 GAGGCCTGTGGTGGTGGTGGTGG 421
```

RESULT 47
AA460713

LOCUS	AA460713 526 bp mRNA linear EST 09-JUN-1997
DEFINITION	zx69f03.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:796733 3' similar to contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION	AA460713
VERSION	AA460713.1 GI:2185833
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 526)
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE	WashU-Merck EST Project 1997
JOURNAL	Unpublished (1997)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: polyT not found

```
           Seq primer: -41m13 fwd. ET from Amersham
           High quality sequence stop: 508.
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:6041207"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:796733"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             1.4%;  Score 23;  DB 1;  Length 526;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        375 GAGGCCTGTGGTGGTGGTGGTGG 397


RESULT 48
BQ012983/c
LOCUS       BQ012983                535 bp    mRNA    linear   EST 26-MAR-2002
DEFINITION  UI-1-BC1p-ayj-e-01-0-UI.s1 NCI_CGAP_Pl3 Homo sapiens cDNA clone
            UI-1-BC1p-ayj-e-01-0-UI 3', mRNA sequence.
ACCESSION   BQ012983
VERSION     BQ012983.1  GI:19737884
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Steven Brown
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 410-436, >(TGG)n#Simple_repeat (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-1-BC1p-ayj-e-01-0-UI"
                     /tissue_type="Placenta"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Pl3"
                     /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_Pl3 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are GA, AGGAA. For
                     additional information, contact: Bento Soares,
                     bento-soares@uiowa.edu
                     TAG_TISSUE=placenta human 8 week
                     TAG_LIB=UI-1-BC1p
                     TAG_SEQ=GA"
ORIGIN

  Query Match             1.4%;  Score 23;  DB 3;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GAGGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||||
Db        443 GAGGCCTGTGGTGGTGGTGGTGG 421


RESULT 49
CF565539
LOCUS       CF565539                641 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION  1115101B01.y1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 5', mRNA
            sequence.
ACCESSION   CF565539
VERSION     CF565539.1  GI:34991622
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
            Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Merchant,S., Quinn,J. and Shrager,J.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1115
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 641.
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant(1998) Methods in
                     Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
```

micromolar Fe) TAP medium (see La Fontaine S, Quinn JM, Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J, Merchant S.LaFontaine et al. (2002) Eukaryotic Cell, 1:736-757) in a shaking illuminated incubator (same conditions as above). CC-1690 cells were grown to mid-log phase in TAP medium in a shaking illuminated incubator to a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089); CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified from each sample, pooled and cDNA synthesized (see Shrager et al, Plant Physiol. 131, 401-408 for details). The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 1.4%; Score 23; DB 6; Length 641;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        248 GCTCTGGTGGTGGCTCCAGTGGT 270
              |||||||||||||||||||||||
Db        579 GCTCTGGTGGTGGCTCCAGTGGT 601
```

RESULT 50
CF559435
LOCUS CF559435 643 bp mRNA linear EST 23-SEP-2003
DEFINITION 1115047A01.y1 C. reinhardtii CC-1690 (mt+), Stress III (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 5', mRNA sequence.
ACCESSION CF559435
VERSION CF559435.1 GI:34985518
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 643)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Merchant,S., Quinn,J. and Shrager,J.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1115
JOURNAL Unpublished (2003)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
High quality sequence stop: 643.
FEATURES Location/Qualifiers
source 1. .643
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690 (mt+), Stress III (normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition III library, constructed by Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from CC-1690 cells grown to mid-log phase in copper-free TAP medium (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated (about 100 micromole/m2/sec) incubator at 22 C (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279); CC-1690 cells grown to mid-log phase in low Fe (1 micromolar Fe) TAP medium (see La Fontaine S, Quinn JM, Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J, Merchant S.LaFontaine et al. (2002) Eukaryotic Cell, 1:736-757) in a shaking illuminated incubator (same conditions as above). CC-1690 cells were grown to mid-log phase in TAP medium in a shaking illuminated incubator to a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089); CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified from each sample, pooled and cDNA synthesized (see Shrager et al, Plant Physiol. 131, 401-408 for details). The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 1.4%; Score 23; DB 6; Length 643;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        248 GCTCTGGTGGTGGCTCCAGTGGT 270
              |||||||||||||||||||||||
Db        579 GCTCTGGTGGTGGCTCCAGTGGT 601
```

Search completed: March 22, 2006, 07:03:41
Job time : 6683 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:30:25 ; Search time 1000 Seconds
(without alignments)
10596.861 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1488	93.6	1590	2	AAV63614	Aav63614 DNA seque
2	1488	93.6	1590	5	AAS43749	Aas43749 DNA encod
3	1488	93.6	1725	9	ABT43850	Abt43850 DNA of th
4	1488	93.6	1725	12	ADN04225	Adn04225 Antipsori
5	1437	90.4	2057	12	ADP10483	Adp10483 Reference
6	1434	90.2	1590	2	AAV63541	Aav63541 DNA seque
7	1355	85.2	4740	9	ABT43849	Abt43849 Partial D
8	1355	85.2	5337	4	AAH45309	Aah45309 Human MHC
9	1355	85.2	25235	4	AAH45310	Aah45310 Human SEE
10	1355	85.2	55050	6	ABQ75680	Abq75680 Human SEE
11	1353	85.1	1506	9	ABT43852	Abt43852 Exon 2 of
12	1245	78.3	2530	6	ABL68811	Abl68811 Kidney ca
13	1245	78.3	2530	9	ABT43848	Abt43848 DNA of th
14	1245	78.3	2530	12	ADQ21503	Adq21503 Human sof
15	978	61.5	2284	3	AAZ50582	Aaz50582 Human epi
16	966	60.8	1725	4	AAH73624	Aah73624 DNA seque
17	933	58.7	2551	12	ADQ25340	Adq25340 Human sof
18	557	35.0	659	12	ACH92172	Ach92172 Human gen
19	505	31.8	505	12	ACH78468	Ach78468 Human gen
20	85	5.3	85	9	ABT43851	Abt43851 Exon 1 of
21	79	5.0	79	5	AAS43704	Aas43704 Corneodes
22	50	3.1	50	5	AAS43513	Aas43513 Corneodes
23	50	3.1	85	5	AAS43706	Aas43706 Corneodes
24	46	2.9	81	5	AAS43709	Aas43709 Corneodes
25	45	2.8	85	5	AAS43705	Aas43705 Corneodes
26	43	2.7	75	5	AAS43710	Aas43710 Corneodes
27	43	2.7	75	6	ABQ75643	Abq75643 Human SEE
28	42	2.6	80	5	AAS43703	Aas43703 Corneodes
29	41	2.6	76	5	AAS43711	Aas43711 Corneodes
30	41	2.6	76	6	ABQ75642	Abq75642 Human SEE
31	41	2.6	79	5	AAS43702	Aas43702 Corneodes
32	41	2.6	81	6	ABQ75645	Abq75645 Human SEE
33	40	2.5	79	5	AAS43708	Aas43708 Corneodes
34	40	2.5	81	5	AAS43700	Aas43700 Corneodes
35	39	2.5	72	5	AAS43707	Aas43707 Corneodes
36	38	2.4	71	6	ABQ75644	Abq75644 Human SEE
37	38	2.4	72	5	AAS43716	Aas43716 Corneodes
38	38	2.4	74	5	AAS43715	Aas43715 Corneodes
39	38	2.4	74	6	ABQ75639	Abq75639 Human SEE
40	37	2.3	65	5	AAS43701	Aas43701 Corneodes
41	36	2.3	71	5	AAS43686	Aas43686 Corneodes
42	35	2.2	69	5	AAS43712	Aas43712 Corneodes
43	33	2.1	67	5	AAS43714	Aas43714 Corneodes
44	33	2.1	67	6	ABQ75640	Abq75640 Human SEE
45	33	2.1	68	5	AAS43685	Aas43685 Corneodes
46	33	2.1	72	6	ABQ75638	Abq75638 Human SEE
47	32	2.0	71	5	AAS43718	Aas43718 Corneodes
48	32	2.0	71	6	ABQ75637	Abq75637 Human SEE
49	31	1.9	66	5	AAS43717	Aas43717 Corneodes
50	30	1.9	50	5	AAS43515	Aas43515 Corneodes
51	29	1.8	50	5	AAS43519	Aas43519 Corneodes
52	29	1.8	50	5	AAS43525	Aas43525 Corneodes
c 53	28	1.8	28	12	ADP12221	Adp12221 Taqman pr
54	28	1.8	30	5	AAS43568	Aas43568 Corneodes
c 55	28	1.8	30	5	AAS43558	Aas43558 Corneodes
56	28	1.8	30	5	AAS43572	Aas43572 Corneodes
57	28	1.8	50	5	AAS43522	Aas43522 Corneodes
58	28	1.8	50	5	AAS43517	Aas43517 Corneodes
59	28	1.8	50	5	AAS43509	Aas43509 Corneodes
60	28	1.8	50	5	AAS43713	Aas43713 Corneodes
61	27	1.7	30	5	AAS43576	Aas43576 Corneodes
c 62	27	1.7	30	5	AAS43554	Aas43554 Corneodes
63	27	1.7	30	5	AAS43574	Aas43574 Corneodes
c 64	27	1.7	30	5	AAS43552	Aas43552 Corneodes
65	27	1.7	50	5	AAS43524	Aas43524 Corneodes
66	27	1.7	50	5	AAS43521	Aas43521 Corneodes
67	27	1.7	50	5	AAS43518	Aas43518 Corneodes
68	26	1.6	26	3	AAC60758	Aac60758 Human pso
69	26	1.6	29	5	AAS43564	Aas43564 Corneodes
70	26	1.6	50	5	AAS43516	Aas43516 Corneodes
71	26	1.6	50	5	AAS43520	Aas43520 Corneodes
72	26	1.6	50	5	AAS43512	Aas43512 Corneodes
73	26	1.6	50	5	AAS43511	Aas43511 Corneodes
74	26	1.6	50	5	AAS43508	Aas43508 Corneodes
75	26	1.6	69	6	ABQ75641	Abq75641 Human SEE

ALIGNMENTS

RESULT 1
AAV63614
ID AAV63614 standard; DNA; 1590 BP.
XX
AC AAV63614;
XX
DT 10-FEB-1999 (first entry)
XX
DE DNA sequence encoding a protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;

KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1590
FT /*tag= a
XX
PN FR2761362-A1.
XX
PD 02-OCT-1998.
XX
PF 28-MAR-1997; 97FR-00003899.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586093/50.
DR P-PSDB; AAW80498.
XX
PT New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT or therapeutic treatment of trophic disorders or thinning of the
PT epidermis.
XX
PS Claim 27; Page 32; 35pp; French.
XX
CC The present sequence encodes a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to
CC reinforce intercorneocyte cohesion and/or to induce thickening of the
CC corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC or benign hyperkeratotic tumours (including those of the Malphigian
CC mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC mouth or during prolapse) can also be treated with the products
XX
SQ Sequence 1590 BP; 291 A; 534 C; 419 G; 346 T; 0 U; 0 Other;

Query Match 93.6%; Score 1488; DB 2; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
```

```
Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
```

RESULT 2
AAS43749
ID AAS43749 standard; DNA; 1590 BP.
XX
AC AAS43749;
XX
DT 18-DEC-2001 (first entry)
XX
DE DNA encoding corneodesmosin.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
DR P-PSDB; AAU27550.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Claim 1; Fig 1; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 1590 BP; 291 A; 534 C; 419 G; 346 T; 0 U; 0 Other;

Query Match 93.6%; Score 1488; DB 5; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
```

```
                |||||||
```

```
Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604


RESULT 4
ADN04225
ID   ADN04225 standard; cDNA; 1725 BP.
XX
AC   ADN04225;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic cDNA sequence #312.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   P-PSDB; ADN04226.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 1; SEQ ID NO 619; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polynucleotides of the invention.
XX
SQ   Sequence 1725 BP; 324 A; 570 C; 454 G; 377 T; 0 U; 0 Other;

  Query Match             93.6%;  Score 1488;  DB 12;  Length 1725;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1588;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
                |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
                |||||||||||||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 5
ADP10483
ID ADP10483 standard; DNA; 2057 BP.
XX
AC ADP10483;
XX
DT 12-AUG-2004 (first entry)
XX
DE Reference mRNA sequences for marker probe #160.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 80; SEQ ID NO 492; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.
XX
SQ Sequence 2057 BP; 416 A; 681 C; 490 G; 470 T; 0 U; 0 Other;

```
  Query Match             90.4%;  Score 1437;  DB 12;  Length 2057;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1587;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194
```

```
Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
            |||||||||||||||||||||||||||||||
Db      1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 6
AAV63541
ID AAV63541 standard; DNA; 1590 BP.
XX
AC AAV63541;
XX
DT 05-FEB-1999 (first entry)
XX
DE DNA sequence encoding a protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1590
FT /*tag= a
XX
PN FR2761363-A1.
XX
PD 02-OCT-1998.
XX
PF 11-SEP-1997; 97FR-00011317.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;
XX
DR WPI; 1998-586094/50.
DR P-PSDB; AAW85055.
XX
PT New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT epidermis etc., also compositions containing protease that degrades
PT corneodesmosin.
XX
PS Claim 27; Page 33; 35pp; French.
XX
CC The present sequence encodes a protein designated corneodesmosin. The
CC protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC or pharmaceutical compositions, to treat trophic disorders of the skin
CC and disorders associated with scar formation, and thinning of the
CC epidermis, particularly the corneal layer, or skin fragility, to

```
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 1590 BP; 291 A; 534 C; 418 G; 347 T; 0 U; 0 Other;

  Query Match             90.2%;  Score 1434;  DB 2;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1534;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTC 114

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGG 174

Qy        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 AAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATT 234

Qy        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCC 294

Qy        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCC 354

Qy        355 TCCGGATCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        355 TCCGGATCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGC 414

Qy        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474

Qy        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 AGCAGCTTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACC 534

Qy        535 AATGACAACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        535 AATGACAACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCT 594

Qy        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 TCCCAAACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGT 654

Qy        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AGTTCGGACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCC 714

Qy        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGAC 774

Qy        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 CAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTT 834

Qy        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 CCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTG 894

Qy        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATC 954

Qy        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 TATCCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCC 1014

Qy       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 TTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCC 1074

Qy       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 AGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAG 1134

Qy       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTC 1194

Qy       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 CCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCT 1254

Qy       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 TCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAA 1314

Qy       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCT 1374

Qy       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 TGCATGTCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGAT 1434

Qy       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCC 1494

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 ATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTT 1554

Qy       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||||||||
Db       1555 TTCCTACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 7
ABT43849
ID   ABT43849 standard; DNA; 4740 BP.
XX
AC   ABT43849;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Partial DNA of the MHC region containing the human corneodesmosin gene.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW   gene; MHC.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            139. .237
FT                   /*tag=  b
FT                   /number= 1
FT                   /note= "This sequence is also found in ABT43851"
FT   CDS             153. .4546
FT                   /*tag=  a
FT                   /product= "Human CDSN protein (corneodesmosin)"
FT                   /transl_except= (pos: 234. .237; aa: Pro)
FT                   /note= "This codon has an apparent one nucleotide
FT                   insertion that alters the reading frame"
```

```
FT                   /transl_except= (pos: 4178. .4180; aa: Ala)
FT                   /note= "This coding sequence contains one intron"
FT   intron          238. .3040
FT                   /*tag=  c
FT                   /number= 1
FT   exon            3041. .4546
FT                   /*tag=  d
FT                   /number= 2
FT                   /note= "This sequence is also found in ABT43852"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   P-PSDB; ABJ39148.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2c; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polynucleotide sequence is the partial sequence of the MHC region that
CC   contains 2 exons coding for the human CDSN gene of the invention
XX
SQ   Sequence 4740 BP; 1050 A; 1305 C; 1207 G; 1178 T; 0 U; 0 Other;

  Query Match             85.2%;  Score 1355;  DB 9;  Length 4740;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3039 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3098

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3099 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3158

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3159 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3218

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3219 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3278

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3279 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3338

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3339 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3398

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3399 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3458

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       3459 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3518

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3519 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3578

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3579 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3638

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3639 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3698

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3699 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3758

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3759 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3818

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3819 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3878

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3879 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3938

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3939 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3998

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3999 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4058

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4059 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4118

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4119 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4178

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4179 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4238

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4239 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4298

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4299 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4358

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4359 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4418

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4419 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4478

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4479 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4538

Qy       1583 GTCCATAA 1590
                ||||||||
Db       4539 GTCCATAA 4546


RESULT 8
AAH45309
ID   AAH45309 standard; DNA; 5337 BP.
XX
AC   AAH45309;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human MHC S DNA.
XX
KW   Human; MHC S; major histocompatibility complex S; vulgar psoriasis;
KW   diagnosis; primer; SEEK1; HCR; a-helix coiled-coil rod homologue;
KW   polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1. .99
FT                   /*tag= a
FT   intron          100. .2903
FT                   /*tag= b
FT   exon            2904. .5337
FT                   /*tag= c
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Claim 1; Page 41-44; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is the human major
CC   histocompatibility complex (MHC) S gene, which was used to design primers
CC   for use in the invention
XX
SQ   Sequence 5337 BP; 1184 A; 1487 C; 1315 G; 1351 T; 0 U; 0 Other;

  Query Match             85.2%;  Score 1355;  DB 4;  Length 5337;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040
```

```
Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4401 GTCCATAA 4408
```

```
RESULT 9
AAH45310
ID   AAH45310 standard; DNA; 25235 BP.
XX
AC   AAH45310;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human SEEK1 DNA.
XX
KW   Human; MHC S; major histocompatibility complex S; vulgar psoriasis;
KW   diagnosis; primer; SEEK1; HCR; a-helix coiled-coil rod homologue;
KW   polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1. .420
FT                   /*tag=  a
FT   intron          421. .1281
FT                   /*tag=  b
FT   exon            1282. .1405
FT                   /*tag=  c
FT   intron          1406. .1601
FT                   /*tag=  d
FT   exon            1602. .1702
FT                   /*tag=  e
FT   intron          1703. .2351
FT                   /*tag=  f
FT   exon            2352. .2364
FT                   /*tag=  g
FT   intron          2365. .6286
FT                   /*tag=  h
FT   exon            6287. .6509
FT                   /*tag=  i
FT   intron          6510. .10416
FT                   /*tag=  j
FT   exon            10417. .10493
FT                   /*tag=  k
FT   intron          10494. .14243
FT                   /*tag=  l
FT   exon            14244. .14407
FT                   /*tag=  m
FT   intron          14408. .14243
FT                   /*tag=  n
FT   exon            25190. .25235
FT                   /*tag=  o
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Claim 1; Page 46-62; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is the human SEEK1 gene
CC   which was used to design primers for use in the invention
XX
SQ   Sequence 25235 BP; 5967 A; 6703 C; 6487 G; 6078 T; 0 U; 0 Other;
```

```
  Query Match             85.2%;  Score 1355;  DB 4;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091
```

```
Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
              ||||||||
Db      24052 GTCCATAA 24059


RESULT 10
ABQ75680
ID   ABQ75680 standard; DNA; 55050 BP.
XX
AC   ABQ75680;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Human SEEK1 consensus genomic DNA.
XX
KW   Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW   antipsoriatic; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200264831-A2.
XX
PD   22-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-GB000653.
XX
PR   13-FEB-2001; 2001GB-00003514.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Lench NJ,  Allen MJ,  Nicholls RK;
XX
DR   WPI; 2002-667013/71.
XX
PT   Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT   gene, for treating, diagnosing or determining susceptibility to SEEK1
PT   mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS   Claim 4; Fig 2; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 nucleotide sequence
XX
SQ   Sequence 55050 BP; 13076 A; 14198 C; 14560 G; 13197 T; 0 U; 19 Other;

  Query Match             85.2%;  Score 1355;  DB 6;  Length 55050;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47898 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 47957

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      47958 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 48017

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48018 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 48077

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48078 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 48137

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      48138 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 48197

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48198 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 48257
```

```
Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48258 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 48317

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     48318 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 48377

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48378 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 48437

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48438 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 48497

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48498 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 48557

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48558 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 48617

Qy       803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48618 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 48677

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48678 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 48737

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48738 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 48797

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48798 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 48857

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48858 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 48917

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48918 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 48977

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48978 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 49037

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49038 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 49097

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49098 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 49157

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49158 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 49217

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49218 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 49277

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49278 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 49337

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49338 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 49397

Qy      1583 GTCCATAA 1590
            ||||||||
Db     49398 GTCCATAA 49405
```

RESULT 11
ABT43852
ID ABT43852 standard; DNA; 1506 BP.
XX
AC ABT43852;
XX
DT 06-NOV-2003 (first entry)
XX
DE Exon 2 of the human corneodesmosin (CDSN) gene.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW gene.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR P-PSDB; ABJ39148.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2c; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polynucleotide sequence is exon 2 of the human corneodesmosin (CDSN) gene
CC of the invention
XX
SQ Sequence 1506 BP; 283 A; 511 C; 384 G; 328 T; 0 U; 0 Other;

```
  Query Match             85.1%;  Score 1353;  DB 9;  Length 1506;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1503;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         85 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 60
```

```
Qy       145 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 204
Db        61 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 120

Qy       205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 264
Db       121 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 180

Qy       265 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 324
Db       181 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 240

Qy       325 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCA 384
Db       241 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCA 300

Qy       385 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 444
Db       301 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 360

Qy       445 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
Db       361 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 420

Qy       505 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAAC 564
Db       421 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAAC 480

Qy       565 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 624
Db       481 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 540

Qy       625 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 684
Db       541 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 600

Qy       685 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
Db       601 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 660

Qy       745 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 804
Db       661 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 720

Qy       805 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 864
Db       721 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 780

Qy       865 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 924
Db       781 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 840

Qy       925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984
Db       841 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 900

Qy       985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
Db       901 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 960

Qy      1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
Db       961 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1020

Qy      1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
Db      1021 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1080

Qy      1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
Db      1081 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCTCC 1140

Qy      1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
Db      1141 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1200

Qy      1285 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
Db      1201 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1260

Qy      1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
Db      1261 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1320

Qy      1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
Db      1321 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1380

Qy      1465 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1524
Db      1381 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1440

Qy      1525 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1584
Db      1441 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1500

Qy      1585 CCATAA 1590
Db      1501 CCATAA 1506
```

RESULT 12
ABL68811
ID ABL68811 standard; DNA; 2530 BP.
XX
AC ABL68811;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:7148.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 7148; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2530 BP; 496 A; 840 C; 593 G; 601 T; 0 U; 0 Other;

Query Match 78.3%; Score 1245; DB 6; Length 2530;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334
```

```
Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

RESULT 13
ABT43848
ID ABT43848 standard; DNA; 2530 BP.
XX
AC ABT43848;
XX
DT 06-NOV-2003 (first entry)
XX
DE DNA of the human S protein (skin corneodesmosin) mRNA.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW gene; S protein; skin.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   63. .1523
FT   /*tag= a
FT   /product= "Human S protein (skin corneodesmosin)"
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR P-PSDB; ABJ39147.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2a; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polynucleotide sequence is the DNA copy of the mRNA that encodes the
CC human S protein (the corneodesmosin protein which is only expressed in
CC the skin) of the invention
XX
SQ Sequence 2530 BP; 496 A; 840 C; 593 G; 601 T; 0 U; 0 Other;

  Query Match   78.3%; Score 1245; DB 9; Length 2530;
  Best Local Similarity 99.7%; Pred. No. 0;
  Matches 1495; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
```

```
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

```
RESULT 14
ADQ21503
ID   ADQ21503 standard; DNA; 2530 BP.
XX
AC   ADQ21503;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 4323.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 4323; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 2530 BP; 496 A; 840 C; 593 G; 601 T; 0 U; 0 Other;

  Query Match             78.3%;  Score 1245;  DB 12;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
```

```
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514


RESULT 15
AAZ50582
ID   AAZ50582 standard; cDNA; 2284 BP.
XX
AC   AAZ50582;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human epidermal protein-6 cDNA.
XX
KW   Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW   dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW   arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1. .142
FT                   /*tag=  b
FT   CDS             47. .1636
FT                   /*tag=  a
FT                   /product= "Human epidermal protein-6"
FT   mat_peptide     148. .1636
FT                   /*tag=  c
FT                   /product= "Mature human epidermal protein-6"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   P-PSDB; AAY44989.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 7; Page 78-79; 82pp; English.
XX
CC   The present cDNA sequence encodes human epidermal protein-6 (HEPI). The
CC   cDNA clone is derived from PTHYNOT03 library which was constructed using
CC   RNA isolated from left parathyroid tissue of a 69-year-old caucasian
CC   female during a partial parathyroidectomy. Recombinant vectors comprising
CC   HEPI cDNA are introduced into host cells for protein expression. The HEPI
CC   proteins are useful for the treatment of epithelial disorders, including
CC   dyshidrotic eczema and scabies, cell proliferative disorders including
CC   actinic keratosis and arteriosclerosis, and autoimmune/inflammatory
CC   disorders like acquired immune deficiency syndrome (AIDS) and Addison's
CC   disease. Pharmaceutical compositions comprising HEPI proteins are useful
CC   for treating disorders associated with altered HEPI expression
XX
SQ   Sequence 2284 BP; 452 A; 775 C; 534 G; 523 T; 0 U; 0 Other;

  Query Match             61.5%;  Score 978;  DB 3;  Length 2284;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1578;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
```

```
Db        467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Db        527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Db        587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Db        647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Db        707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Db        767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Db        827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Db        887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Db        947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Db       1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Db       1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Db       1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Db       1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Db       1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Db       1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Db       1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Db       1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Db       1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Db       1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590

Db       1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636
```

RESULT 16
AAH73624
ID AAH73624 standard; DNA; 1725 BP.
XX
AC AAH73624;
XX
DT 01-OCT-2001 (first entry)
XX
DE DNA sequence #1 for vulgar psoriasis diagnosis.
XX
KW MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism; ds.
XX
OS Unidentified.
XX
PN WO200142458-A1.
XX
PD 14-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-JP008624.
XX
PR 06-DEC-1999; 99JP-00346867.
XX
PA (INOK/) INOKO H.
XX
PI Inoko H, Tamiya G;
XX
DR WPI; 2001-381680/40.
DR P-PSDB; AAG64152.
XX
PT New primer DNA, useful for detecting vulgar psoriasis.
XX
PS Disclosure; Fig 2; 106pp; Japanese.
XX
CC The invention relates to a method of diagnosing vulgar psoriasis using
CC primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC By analysing the sequences of these genes in Japanese patients with
CC psoriasis and in normal subjects, it has been found that some of the
CC examined polymorphisms correlate significantly to the group of patients
CC with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC these gene polymorphisms. The present sequence is provided in the
CC specification
XX
SQ Sequence 1725 BP; 322 A; 566 C; 449 G; 377 T; 0 U; 11 Other;

```
  Query Match            60.8%;  Score 966;  DB 4;  Length 1725;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1566;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
```

```
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Db        435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Db        615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Db        855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Db       1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTC 1578

Db       1575 CCCCAAGGAGAGTTACTC 1592
```

```
RESULT 17
ADQ25340
ID   ADQ25340 standard; DNA; 2551 BP.
XX
AC   ADQ25340;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 8160.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 8160; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 2551 BP; 471 A; 789 C; 559 G; 563 T; 0 U; 169 Other;

  Query Match           58.7%;  Score 933;  DB 12;  Length 2551;
  Best Local Similarity 99.9%;  Pred. No. 0;
```

```
  Matches  983;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 637

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 697

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 757

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 817

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 877

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 937

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 997

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1057

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1117

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1118 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1177

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1178 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1237

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1238 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1297

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1298 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1357

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1358 ATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1417

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1418 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1477

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1478 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1537

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCC 1533
              |||||||||||||||||||||||||
Db       1538 GCCCAAGTGAAGCCTCTGGGGCCC 1561


RESULT 18
ACH92172
ID   ACH92172 standard; DNA; 659 BP.
XX
AC   ACH92172;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #25367.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 25367; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 659 BP; 132 A; 222 C; 166 G; 139 T; 0 U; 0 Other;
```

```
  Query Match             35.0%;  Score 557;  DB 12;  Length 659;
  Best Local Similarity   99.7%;  Pred. No. 3.9e-241;
  Matches  657;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         86 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 60

Qy        146 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 120

Qy        206 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 180

Qy        266 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAG 240

Qy        326 GAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCAT 385
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 GAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCAT 300

Qy        386 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGG 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGG 360

Qy        446 GAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 420

Qy        506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        421 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAACC 480

Qy        566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 540

Qy        626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 600

Qy        686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 659
```

```
RESULT 19
ACH78468
ID   ACH78468 standard; DNA; 505 BP.
XX
AC   ACH78468;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #11663.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 15; SEQ ID NO 11663; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 505 BP; 89 A; 165 C; 138 G; 113 T; 0 U; 0 Other;
```

```
  Query Match             31.8%;  Score 505;  DB 12;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-217;
  Matches  505;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        629 GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAG 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAG 60

Qy        689 GGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTC 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTC 120

Qy        749 AGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAG 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAG 180

Qy        809 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 240
```

```
Qy        869 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 300

Qy        929 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 360

Qy        989 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 420

Qy       1049 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 480

Qy       1109 AGCCAGTGGGGACTGGTGGGGTCCA 1133
              |||||||||||||||||||||||||
Db        481 AGCCAGTGGGGACTGGTGGGGTCCA 505
```

RESULT 20
ABT43851
ID ABT43851 standard; DNA; 85 BP.
XX
AC ABT43851;
XX
DT 06-NOV-2003 (first entry)
XX
DE Exon 1 of the human corneodesmosin (CDSN) gene.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness; ds;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW gene.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR P-PSDB; ABJ39148.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2c; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polynucleotide sequence is exon 1 of the human corneodesmosin (CDSN) gene
CC of the invention
XX
SQ Sequence 85 BP; 8 A; 23 C; 35 G; 19 T; 0 U; 0 Other;

```
  Query Match              5.3%;  Score 85;  DB 9;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-28;
  Matches   85;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAG 85
              |||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAG 85
```

RESULT 21
AAS43704
ID AAS43704 standard; DNA; 79 BP.
XX
AC AAS43704;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #62.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention

```
XX
SQ   Sequence 79 BP; 18 A; 25 C; 24 G; 12 T; 0 U; 0 Other;

  Query Match             5.0%;  Score 79;  DB 5;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-25;
  Matches   79;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 GGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATT 463
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATT 60

Qy        464 CGAGCAGCAGCAGCAGCTT 482
              |||||||||||||||||||
Db         61 CGAGCAGCAGCAGCAGCTT 79


RESULT 22
AAS43513
ID   AAS43513 standard; DNA; 50 BP.
XX
AC   AAS43513;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #10.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 35; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 50 BP; 12 A; 16 C; 15 G; 7 T; 0 U; 0 Other;

  Query Match             3.1%;  Score 50;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        425 GAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 50


RESULT 23
AAS43706
ID   AAS43706 standard; DNA; 85 BP.
XX
AC   AAS43706;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #64.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 85 BP; 21 A; 30 C; 15 G; 18 T; 0 U; 1 Other;

  Query Match             3.1%;  Score 50;  DB 5;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-12;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCT 604
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCT 50


RESULT 24
AAS43709
ID   AAS43709 standard; DNA; 81 BP.
XX
AC   AAS43709;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #67.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 81 BP; 24 A; 19 C; 18 G; 19 T; 0 U; 1 Other;

  Query Match             2.9%;  Score 46;  DB 5;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-10;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        911 ACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTA 956
              ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTA 46


RESULT 25
AAS43705
ID   AAS43705 standard; DNA; 85 BP.
XX
AC   AAS43705;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #63.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 85 BP; 20 A; 30 C; 15 G; 18 T; 0 U; 2 Other;

  Query Match             2.8%;  Score 45;  DB 5;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-10;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 599
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 45


RESULT 26
AAS43710
ID   AAS43710 standard; DNA; 75 BP.
XX
AC   AAS43710;
XX
DT   18-DEC-2001  (first entry)
XX
```

DE Corneodesmosin single nucleotide polymorphism #68.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 75 BP; 11 A; 28 C; 21 G; 14 T; 0 U; 1 Other;

```
  Query Match             2.7%;  Score 43;  DB 5;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 7e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1061 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 1103
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 43
```

RESULT 27
ABQ75643
ID ABQ75643 standard; DNA; 75 BP.
XX
AC ABQ75643;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human SEEK1 intron 1 SNP oligonucleotide #8.
XX
KW Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS Homo sapiens.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.
XX
PT Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT gene, for treating, diagnosing or determining susceptibility to SEEK1
PT mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS Example; Page 35; 69pp; English.
XX
CC The present invention describes a polynucleotide (I) comprising: (a) a
CC nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC acid sequence (IV) that hybridises under stringent conditions to the
CC polynucleotide of (II) or (III). (I) has antiinflammatory and
CC antipsoriatic activities and can be used in gene therapy. The alleles of
CC one or more polymorphisms in the SEEK1 gene can be used for the
CC manufacture of a medicament for the diagnosis and treatment of SEEK1
CC mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC prognostic or therapeutic methods, or as research tools in drug
CC screening. Human SEEK1 is located to chromosome 6p21. The present
CC sequence represents a human SEEK1 oligonucleotide comprising a single
CC nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC ambiguity base
XX
SQ Sequence 75 BP; 11 A; 28 C; 21 G; 14 T; 0 U; 1 Other;

```
  Query Match             2.7%;  Score 43;  DB 6;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 7e-09;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1061 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 1103
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 43
```

RESULT 28
AAS43703
ID AAS43703 standard; DNA; 80 BP.
XX
AC AAS43703;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #61.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX

PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 80 BP; 14 A; 28 C; 23 G; 14 T; 0 U; 1 Other;

Query Match 2.6%; Score 42; DB 5; Length 80;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 386 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAA 427
||||||||||||||||||||||||||||||||||||||||||
Db 1 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAA 42

RESULT 29
AAS43711
ID AAS43711 standard; DNA; 76 BP.
XX
AC AAS43711;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #69.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 76 BP; 12 A; 31 C; 14 G; 18 T; 0 U; 1 Other;

Query Match 2.6%; Score 41; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
|||||||||||||||||||||||||||||||||||||||||
Db 1 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 41

RESULT 30
ABQ75642
ID ABQ75642 standard; DNA; 76 BP.
XX
AC ABQ75642;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human SEEK1 intron 1 SNP oligonucleotide #7.
XX
KW Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.
XX
PT Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT gene, for treating, diagnosing or determining susceptibility to SEEK1
PT mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS Example; Page 34; 69pp; English.
XX
CC The present invention describes a polynucleotide (I) comprising: (a) a

```
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 76 BP; 12 A; 30 C; 14 G; 17 T; 0 U; 3 Other;

  Query Match             2.6%;  Score 41;  DB 6;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-08;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              |||||||||||||||||||||||||||||||||||||||||
Db          1 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 41


RESULT 31
AAS43702
ID   AAS43702 standard; DNA; 79 BP.
XX
AC   AAS43702;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #60.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 79 BP; 15 A; 25 C; 22 G; 16 T; 0 U; 1 Other;

  Query Match             2.6%;  Score 41;  DB 5;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-08;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 233
              |||||||||||||||||||||||||||||||||||||||||
Db         39 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 79


RESULT 32
ABQ75645
ID   ABQ75645 standard; DNA; 81 BP.
XX
AC   ABQ75645;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Human SEEK1 intron 1 SNP oligonucleotide #10.
XX
KW   Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW   antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200264831-A2.
XX
PD   22-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-GB000653.
XX
PR   13-FEB-2001; 2001GB-00003514.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Lench NJ,  Allen MJ,  Nicholls RK;
XX
DR   WPI; 2002-667013/71.
XX
PT   Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT   gene, for treating, diagnosing or determining susceptibility to SEEK1
PT   mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS   Example; Page 35; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 81 BP; 18 A; 31 C; 16 G; 15 T; 0 U; 1 Other;

  Query Match             2.6%;  Score 41;  DB 6;  Length 81;
```

```
  Best Local Similarity   100.0%;  Pred. No. 5.6e-08;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        123 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCT 163
              |||||||||||||||||||||||||||||||||||||||||
Db         41 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCT 81


RESULT 33
AAS43708'
ID   AAS43708 standard; DNA; 79 BP.
XX
AC   AAS43708;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #66.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 79 BP; 10 A; 23 C; 27 G; 18 T; 0 U; 1 Other;

  Query Match              2.5%;  Score 40;  DB 5;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-07;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        713 CCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAG 752
              ||||||||||||||||||||||||||||||||||||||||
Db          1 CCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAG 40


RESULT 34
AAS43700
ID   AAS43700 standard; DNA; 81 BP.
XX
AC   AAS43700;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #58.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 81 BP; 18 A; 31 C; 16 G; 15 T; 0 U; 1 Other;

  Query Match              2.5%;  Score 40;  DB 5;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-07;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 122
              ||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 40


RESULT 35
AAS43707
ID   AAS43707 standard; DNA; 72 BP.
XX
```

```
AC   AAS43707;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #65.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 72 BP; 9 A; 32 C; 15 G; 15 T; 0 U; 1 Other;

  Query Match             2.5%;  Score 39;  DB 5;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-07;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        669 CGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTC 707
              |||||||||||||||||||||||||||||||||||||||
Db          1 CGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTC 39


RESULT 36
ABQ75644
ID   ABQ75644 standard; DNA; 71 BP.
XX
AC   ABQ75644;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Human SEEK1 intron 1 SNP oligonucleotide #9.
XX
KW   Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW   antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200264831-A2.
XX
PD   22-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-GB000653.
XX
PR   13-FEB-2001; 2001GB-00003514.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Lench NJ,  Allen MJ,  Nicholls RK;
XX
DR   WPI; 2002-667013/71.
XX
PT   Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT   gene, for treating, diagnosing or determining susceptibility to SEEK1
PT   mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS   Example; Page 35; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 71 BP; 23 A; 17 C; 17 G; 13 T; 0 U; 1 Other;

  Query Match             2.4%;  Score 38;  DB 6;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGG 995
              |||||||||||||||||||||||||||||||||||||
Db         34 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGG 71


RESULT 37
AAS43716
ID   AAS43716 standard; DNA; 72 BP.
XX
AC   AAS43716;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #74.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
```

PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 40; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 72 BP; 15 A; 23 C; 15 G; 18 T; 0 U; 1 Other;

Query Match 2.4%; Score 38; DB 5; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1306 AGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTG 1343
||||||||||||||||||||||||||||||||||||||
Db 1 AGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTG 38

RESULT 38
AAS43715
ID AAS43715 standard; DNA; 74 BP.
XX
AC AAS43715;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #73.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 40; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 74 BP; 17 A; 25 C; 16 G; 15 T; 0 U; 1 Other;

Query Match 2.4%; Score 38; DB 5; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 1316
||||||||||||||||||||||||||||||||||||||
Db 1 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 38

RESULT 39
ABQ75639
ID ABQ75639 standard; DNA; 74 BP.
XX
AC ABQ75639;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human SEEK1 intron 1 SNP oligonucleotide #4.
XX
KW Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.
XX
PT Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT gene, for treating, diagnosing or determining susceptibility to SEEK1
PT mediated disease, e.g. inflammatory disease such as psoriasis.

```
XX
PS   Example; Page 34; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 74 BP; 17 A; 25 C; 16 G; 14 T; 0 U; 2 Other;

  Query Match             2.4%;  Score 38;  DB 6;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 1316
              |||||||||||||||||||||||||||||||||||||||
Db          1 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 38


RESULT 40
AAS43701
ID   AAS43701 standard; DNA; 65 BP.
XX
AC   AAS43701;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #59.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 65 BP; 12 A; 26 C; 16 G; 10 T; 0 U; 1 Other;

  Query Match             2.3%;  Score 37;  DB 5;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-06;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       129 GGACCCCACGCGTATCACCTCCCCTAACGACCCCTGC 165
             |||||||||||||||||||||||||||||||||||||
Db         1 GGACCCCACGCGTATCACCTCCCCTAACGACCCCTGC 37


RESULT 41
AAS43686
ID   AAS43686 standard; DNA; 71 BP.
XX
AC   AAS43686;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #44.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
```

CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 71 BP; 6 A; 19 C; 30 G; 15 T; 0 U; 1 Other;

Query Match 2.3%; Score 36; DB 5; Length 71;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 GCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATG 51
||||||||||||||||||||||||||||||||||||
Db 1 GCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATG 36

RESULT 42
AAS43712
ID AAS43712 standard; DNA; 69 BP.
XX
AC AAS43712;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #70.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 69 BP; 11 A; 27 C; 14 G; 16 T; 0 U; 1 Other;

Query Match 2.2%; Score 35; DB 5; Length 69;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 CTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGC 1221
|||||||||||||||||||||||||||||||||||
Db 1 CTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGC 35

RESULT 43
AAS43714
ID AAS43714 standard; DNA; 67 BP.
XX
AC AAS43714;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #72.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 40; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 67 BP; 12 A; 25 C; 14 G; 15 T; 0 U; 1 Other;

Query Match 2.1%; Score 33; DB 5; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1230 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 1262
              ||||||||||||||||||||||||||||||||||
Db         35 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 67


RESULT 44
ABQ75640
ID   ABQ75640 standard; DNA; 67 BP.
XX
AC   ABQ75640;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Human SEEK1 intron 1 SNP oligonucleotide #5.
XX
KW   Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW   antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200264831-A2.
XX
PD   22-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-GB000653.
XX
PR   13-FEB-2001; 2001GB-00003514.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Lench NJ,  Allen MJ,  Nicholls RK;
XX
DR   WPI; 2002-667013/71.
XX
PT   Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT   gene, for treating, diagnosing or determining susceptibility to SEEK1
PT   mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS   Example; Page 34; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 67 BP; 11 A; 25 C; 14 G; 14 T; 0 U; 3 Other;

  Query Match             2.1%;  Score 33;  DB 6;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1230 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 1262
              ||||||||||||||||||||||||||||||||||
Db         35 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 67


RESULT 45
AAS43685
ID   AAS43685 standard; DNA; 68 BP.
XX
AC   AAS43685;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #43.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 39; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 68 BP; 10 A; 19 C; 27 G; 11 T; 0 U; 1 Other;

  Query Match             2.1%;  Score 33;  DB 5;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGT 33
              |||||||||||||||||||||||||||||||||
Db         36 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGT 68


RESULT 46
ABQ75638
ID   ABQ75638 standard; DNA; 72 BP.
XX
AC   ABQ75638;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Human SEEK1 intron 1 SNP oligonucleotide #3.
XX
KW   Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW   antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
```

XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.
XX
PT Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT gene, for treating, diagnosing or determining susceptibility to SEEK1
PT mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS Example; Page 34; 69pp; English.
XX
CC The present invention describes a polynucleotide (I) comprising: (a) a
CC nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC acid sequence (IV) that hybridises under stringent conditions to the
CC polynucleotide of (II) or (III). (I) has antiinflammatory and
CC antipsoriatic activities and can be used in gene therapy. The alleles of
CC one or more polymorphisms in the SEEK1 gene can be used for the
CC manufacture of a medicament for the diagnosis and treatment of SEEK1
CC mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC prognostic or therapeutic methods, or as research tools in drug
CC screening. Human SEEK1 is located to chromosome 6p21. The present
CC sequence represents a human SEEK1 oligonucleotide comprising a single
CC nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC ambiguity base
XX
SQ Sequence 72 BP; 15 A; 23 C; 14 G; 18 T; 0 U; 2 Other;

Query Match 2.1%; Score 33; DB 6; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377
||||||||||||||||||||||||||||||||
Db 40 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 72

RESULT 47
AAS43718
ID AAS43718 standard; DNA; 71 BP.
XX
AC AAS43718;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #76.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB000795.
XX
PR 23-FEB-2000; 2000GB-00004312.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahnini R;
XX
DR WPI; 2001-570627/64.
XX
PT Corneodesmosin protein and polynucleotide encoding it, having one or more
PT polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT diseases.
XX
PS Disclosure; Page 40; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 71 BP; 20 A; 16 C; 18 G; 17 T; 0 U; 0 Other;

Query Match 2.0%; Score 32; DB 5; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1559 TACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
||||||||||||||||||||||||||||||||
Db 1 TACCCCAAGGAGAGTTACTCGACAGTCCATAA 32

RESULT 48
ABQ75637
ID ABQ75637 standard; DNA; 71 BP.
XX
AC ABQ75637;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human SEEK1 intron 1 SNP oligonucleotide #2.
XX
KW Human; SEEK1; chromosome 6p21; inflammatory disease; antiinflammatory;
KW antipsoriatic; gene therapy; SNP; single nucleotide polymorphism; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200264831-A2.
XX
PD 22-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-GB000653.
XX
PR 13-FEB-2001; 2001GB-00003514.
XX
PA (OXAG-) OXAGEN LTD.
XX
PI Lench NJ, Allen MJ, Nicholls RK;
XX
DR WPI; 2002-667013/71.

```
XX
PT   Polymorphisms in the SEEK1 gene and polynucleotides encoding the SEEK1
PT   gene, for treating, diagnosing or determining susceptibility to SEEK1
PT   mediated disease, e.g. inflammatory disease such as psoriasis.
XX
PS   Example; Page 34; 69pp; English.
XX
CC   The present invention describes a polynucleotide (I) comprising: (a) a
CC   nucleic acid sequence (II) encoding the SEEK1 gene; (b) a nucleic acid
CC   sequence (III) encoding a fragment of the SEEK1 gene; or (c) a nucleic
CC   acid sequence (IV) that hybridises under stringent conditions to the
CC   polynucleotide of (II) or (III). (I) has antiinflammatory and
CC   antipsoriatic activities and can be used in gene therapy. The alleles of
CC   one or more polymorphisms in the SEEK1 gene can be used for the
CC   manufacture of a medicament for the diagnosis and treatment of SEEK1
CC   mediated disease. SEEK1 protein fragments are useful in diagnostic,
CC   prognostic or therapeutic methods, or as research tools in drug
CC   screening. Human SEEK1 is located to chromosome 6p21. The present
CC   sequence represents a human SEEK1 oligonucleotide comprising a single
CC   nucleotide polymorphism (SNP), where the SNP is indicated by an IUPAC
CC   ambiguity base
XX
SQ   Sequence 71 BP; 20 A; 16 C; 18 G; 17 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 32;  DB 6;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 0.00065;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1559 TACCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              ||||||||||||||||||||||||||||||||
Db          1 TACCCCAAGGAGAGTTACTCGACAGTCCATAA 32


RESULT 49
AAS43717
ID   AAS43717 standard; DNA; 66 BP.
XX
AC   AAS43717;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #75.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 40; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC   sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC   invention
XX
SQ   Sequence 66 BP; 18 A; 14 C; 14 G; 19 T; 0 U; 1 Other;

  Query Match             1.9%;  Score 31;  DB 5;  Length 66;
  Best Local Similarity   100.0%;  Pred. No. 0.0019;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1548 TGAAGTTTTCCTACCCCAAGGAGAGTTACTC 1578
              |||||||||||||||||||||||||||||||
Db          1 TGAAGTTTTCCTACCCCAAGGAGAGTTACTC 31


RESULT 50
AAS43515
ID   AAS43515 standard; DNA; 50 BP.
XX
AC   AAS43515;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin single nucleotide polymorphism #12.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Disclosure; Page 35; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
```

CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention
XX
SQ Sequence 50 BP; 10 A; 16 C; 11 G; 12 T; 0 U; 1 Other;

Query Match 1.9%; Score 30; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        575 CTGGACAAAGCTCTTCCTCTTCCCAAACCT 604
              ||||||||||||||||||||||||||||||
Db          1 CTGGACAAAGCTCTTCCTCTTCCCAAACCT 30
```

Search completed: March 22, 2006, 02:56:46
Job time : 1006 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 03:02:18 ; Search time 1369 Seconds
(without alignments)
9604.318 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 0

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1539	96.8	1590	9	US-10-677-865-2	Sequence 2, Appli
c	2	1488	93.6	1590	6	US-10-204-884-4	Sequence 4, Appli
	3	1488	93.6	1725	6	US-10-316-175-14	Sequence 14, Appl
	4	1355	85.2	4740	6	US-10-316-175-13	Sequence 13, Appl
	5	1355	85.2	5337	6	US-10-164-230-1	Sequence 1, Appli
	6	1355	85.2	5700	6	US-10-204-884-1	Sequence 1, Appli
	7	1355	85.2	25235	6	US-10-164-230-2	Sequence 2, Appli
	8	1355	85.2	55050	7	US-10-467-752-4	Sequence 4, Appli
	9	1353	85.1	1506	6	US-10-316-175-68	Sequence 68, Appl
	10	1245	78.3	2530	3	US-09-968-007A-678	Sequence 678, App
	11	1245	78.3	2530	6	US-10-316-175-11	Sequence 11, Appl
	12	1245	78.3	2530	8	US-10-723-860-4323	Sequence 4323, Ap
	13	1245	78.3	2530	9	US-10-843-641A-7148	Sequence 7148, Ap
	14	978	61.5	2284	9	US-10-883-805-12	Sequence 12, Appl
	15	966	60.8	1725	6	US-10-164-230-60	Sequence 60, Appl
	16	933	58.7	2551	8	US-10-723-860-8160	Sequence 8160, Ap
	17	621	39.1	1590	6	US-10-204-884-2	Sequence 2, Appli
	18	557	35.0	659	6	US-10-029-386-25367	Sequence 25367, A
	19	505	31.8	505	6	US-10-029-386-11663	Sequence 11663, A
	20	288	18.1	580	4	US-09-925-065A-244963	Sequence 244963,
	21	247	15.5	580	4	US-09-925-065A-244955	Sequence 244955,
	22	237	14.9	580	4	US-09-925-065A-244956	Sequence 244956,
	23	85	5.3	99	6	US-10-316-175-67	Sequence 67, Appl
	24	79	5.0	79	6	US-10-204-884-145	Sequence 145, App
	25	50	3.1	50	6	US-10-204-884-26	Sequence 26, Appl
	26	50	3.1	85	6	US-10-204-884-147	Sequence 147, App
	27	46	2.9	81	6	US-10-204-884-150	Sequence 150, App
	28	45	2.8	85	6	US-10-204-884-146	Sequence 146, App
	29	43	2.7	75	6	US-10-204-884-151	Sequence 151, App
	30	43	2.7	75	7	US-10-467-752-74	Sequence 74, Appl
	31	42	2.6	80	6	US-10-204-884-144	Sequence 144, App
	32	41	2.6	76	6	US-10-204-884-152	Sequence 152, App
	33	41	2.6	76	7	US-10-467-752-73	Sequence 73, Appl
	34	41	2.6	79	6	US-10-204-884-143	Sequence 143, App
	35	41	2.6	81	7	US-10-467-752-76	Sequence 76, Appl
	36	40	2.5	79	6	US-10-204-884-149	Sequence 149, App
	37	40	2.5	81	6	US-10-204-884-141	Sequence 141, App
	38	39	2.5	72	6	US-10-204-884-148	Sequence 148, App
	39	38	2.4	71	7	US-10-467-752-75	Sequence 75, Appl
	40	38	2.4	72	6	US-10-204-884-157	Sequence 157, App
	41	38	2.4	74	6	US-10-204-884-156	Sequence 156, App
	42	38	2.4	74	7	US-10-467-752-70	Sequence 70, Appl
	43	37	2.3	65	6	US-10-204-884-142	Sequence 142, App
	44	36	2.3	71	6	US-10-204-884-127	Sequence 127, App
	45	35	2.2	69	6	US-10-204-884-153	Sequence 153, App
	46	33	2.1	67	6	US-10-204-884-155	Sequence 155, App
	47	33	2.1	67	7	US-10-467-752-71	Sequence 71, Appl
	48	33	2.1	68	6	US-10-204-884-126	Sequence 126, App
	49	33	2.1	72	7	US-10-467-752-69	Sequence 69, Appl
c	50	33	2.1	3673778	6	US-10-312-841-1	Sequence 1, Appli
	51	32	2.0	71	6	US-10-204-884-159	Sequence 159, App
	52	32	2.0	71	7	US-10-467-752-68	Sequence 68, Appl
	53	31	1.9	68	6	US-10-204-884-158	Sequence 158, App
	54	30	1.9	50	6	US-10-204-884-28	Sequence 28, Appl
	55	29	1.8	50	6	US-10-204-884-32	Sequence 32, Appl
	56	29	1.8	50	6	US-10-204-884-38	Sequence 38, Appl
c	57	28	1.8	30	6	US-10-204-884-68	Sequence 68, Appl
	58	28	1.8	30	6	US-10-204-884-78	Sequence 78, Appl
	59	28	1.8	30	6	US-10-204-884-80	Sequence 80, Appl
	60	28	1.8	50	6	US-10-204-884-22	Sequence 22, Appl
	61	28	1.8	50	6	US-10-204-884-30	Sequence 30, Appl
	62	28	1.8	50	6	US-10-204-884-35	Sequence 35, Appl
	63	28	1.8	50	6	US-10-204-884-154	Sequence 154, App
c	64	27	1.7	30	6	US-10-204-884-62	Sequence 62, Appl
c	65	27	1.7	30	6	US-10-204-884-64	Sequence 64, Appl
	66	27	1.7	30	6	US-10-204-884-81	Sequence 81, Appl
	67	27	1.7	30	6	US-10-204-884-83	Sequence 83, Appl
	68	27	1.7	50	6	US-10-204-884-31	Sequence 31, Appl
	69	27	1.7	50	6	US-10-204-884-34	Sequence 34, Appl
	70	27	1.7	50	6	US-10-204-884-37	Sequence 37, Appl
	71	26	1.6	29	6	US-10-204-884-74	Sequence 74, Appl
	72	26	1.6	50	6	US-10-204-884-21	Sequence 21, Appl
	73	26	1.6	50	6	US-10-204-884-24	Sequence 24, Appl
	74	26	1.6	50	6	US-10-204-884-25	Sequence 25, Appl
	75	26	1.6	50	6	US-10-204-884-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-10-677-865-2
; Sequence 2, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/10/677,865
; CURRENT FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: US/09/381,656
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899

```
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1590
;   TYPE: DNA
;   ORGANISM: Human
US-10-677-865-2

  Query Match             96.8%;  Score 1539;  DB 9;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1589;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 2
US-10-204-884-4/c
; Sequence 4, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 1590
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-10-204-884-4

Query Match 93.6%; Score 1488; DB 6; Length 1590;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1590 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 1531

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1530 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 1471

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1470 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 1411

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1410 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 1351

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 1291

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1290 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 1231

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1230 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 1171

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1170 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 1111

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1110 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 1051

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 991

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        990 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 931

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 871

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 811

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 751

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 691

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 631

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 571
```

```
Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 511

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 451

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 391

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 331

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 271

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 211

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 151

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 91

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 31

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db          30 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1
```

RESULT 3
US-10-316-175-14
; Sequence 14, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-316-175-14

Query Match 93.6%; Score 1488; DB 6; Length 1725;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy        61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy       121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy       181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy       241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy       301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy       361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy       421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy       481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy       541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy       601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy       661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy       721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy       781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy       841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy       901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy       961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy      1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy      1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy      1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy      1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1215 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy      1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy      1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy      1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy      1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy      1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy      1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             |||||||||||||||||||||||||||||
Db      1575 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1604
```

RESULT 4
US-10-316-175-13
; Sequence 13, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 4740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-316-175-13

Query Match 85.2%; Score 1355; DB 6; Length 4740;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3039 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3098

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3099 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3158

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3159 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3218

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3219 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3278

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       3279 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3338

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3339 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3398

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3399 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3458

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3459 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3518

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3519 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3578

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3579 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3638

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3639 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3698

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3699 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3758

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3759 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3818

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3819 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3878

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3879 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3938

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3939 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3998

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3999 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4058

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4059 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4118

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4119 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4178

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4179 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4238

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4239 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4298

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4299 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4358

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4359 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4418

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4419 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4478

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4479 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4538

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4539 GTCCATAA 4546
```

RESULT 5
US-10-164-230-1
; Sequence 1, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)...(99)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2904)...(5337)
US-10-164-230-1

Query Match 85.2%; Score 1355; DB 6; Length 5337;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
```

```
Db       3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4401 GTCCATAA 4408
```

RESULT 6
US-10-204-884-1
; Sequence 1, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-1

Query Match 85.2%; Score 1355; DB 6; Length 5700;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3299 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 3358

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3359 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3418

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3419 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3478

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3479 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3538

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3539 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3598

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3599 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3658

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3659 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3718

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3719 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3778
```

```
Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3779 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3838

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3839 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3898

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3899 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3958

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3959 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 4018

Qy       803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4019 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 4078

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4079 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 4138

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4139 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 4198

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4199 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 4258

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4259 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 4318

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4319 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 4378

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4379 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4438

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4439 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4498

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4499 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4558

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4559 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4618

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4619 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4678

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4679 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4738

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4739 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4798

Qy      1583 GTCCATAA 1590
            ||||||||
Db      4799 GTCCATAA 4806
```

```
RESULT 7
US-10-164-230-2
; Sequence 2, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 25235
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(420)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1282)...(1405)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1702)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1631)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (2352)...(2364)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (6287)...(6509)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (10417)...(10493)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14407)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14344)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (25190)...(25235)
US-10-164-230-2

  Query Match             85.2%;  Score 1355;  DB 6;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791
```

```
Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy       443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy       503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy       563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy       623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy       683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy       743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy       803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23272 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy       863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy       923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy       983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy      1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy      1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy      1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy      1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy      1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy      1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy      1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy      1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy      1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy      1583 GTCCATAA 1590
            ||||||||
Db     24052 GTCCATAA 24059
```

RESULT 8
US-10-467-752-4
; Sequence 4, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
; APPLICANT: Lench, et al.
; TITLE OF INVENTION: Test and Model for Inflammatory disease
; FILE REFERENCE: 2003882-0009
; CURRENT APPLICATION NUMBER: US/10/467,752
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: PCT/GB02/00653
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: GB0103514.6
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 55050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13351)..(13351)
; OTHER INFORMATION: n is a or t or g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13401)..(13401)
; OTHER INFORMATION: n is a or t or g or c
US-10-467-752-4

Query Match 85.2%; Score 1355; DB 7; Length 55050;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47898 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 47957

Qy       143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47958 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 48017

Qy       203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48018 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 48077

Qy       263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48078 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 48137

Qy       323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     48138 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 48197

Qy       383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      48198 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 48257

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48258 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 48317

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      48318 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 48377

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48378 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 48437

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48438 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 48497

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48498 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 48557

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48558 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 48617

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48618 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 48677

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48678 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 48737

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48738 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 48797

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48798 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 48857

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48858 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 48917

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48918 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 48977

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48978 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 49037

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49038 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 49097

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49098 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 49157

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49158 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 49217

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49218 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 49277

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49278 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 49337

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49338 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 49397

Qy       1583 GTCCATAA 1590
              ||||||||
Db      49398 GTCCATAA 49405
```

RESULT 9
US-10-316-175-68
; Sequence 68, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-316-175-68

Query Match 85.1%; Score 1353; DB 6; Length 1506;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1503; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          85 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATC 60

Qy         145 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTAC 120

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCC 180

Qy         265 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCA 240

Qy         325 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCA 384
               ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         241 GGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCA 300

Qy         385 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 444
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCG 360

Qy         445 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 504
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTC 420

Qy         505 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAAC 564
               ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         421 CAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAAC 480

Qy         565 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 624
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 CCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGC 540

Qy        625 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGT 600

Qy        685 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 660

Qy        745 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTT 720

Qy        805 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCT 780

Qy        865 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTT 840

Qy        925 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAAC 900

Qy        985 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGC 960

Qy       1045 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCG 1020

Qy       1105 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAG 1080

Qy       1165 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCC 1224
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1081 GGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCTCC 1140

Qy       1225 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACC 1200

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGT 1260

Qy       1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGACT 1320

Qy       1405 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCC 1380

Qy       1465 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCT 1440

Qy       1525 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGT 1500

Qy       1585 CCATAA 1590
              ||||||
Db       1501 CCATAA 1506


RESULT 10
US-09-968-007A-678
```

```
; Sequence 678, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Usining Sign
;  TITLE OF INVENTION:  Gene Sets
;  FILE REFERENCE: 689290-71
;  CURRENT APPLICATION NUMBER: US/09/968,007A
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,172
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,173
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,278
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,294
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,295
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,316
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 1001
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 678
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-968-007A-678

  Query Match             78.3%;  Score 1245;  DB 3;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614
```

```
Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

```
RESULT 11
US-10-316-175-11
; Sequence 11, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 2530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-316-175-11

  Query Match             78.3%;  Score 1245;  DB 6;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854
```

```
Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

RESULT 12
US-10-723-860-4323
; Sequence 4323, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4323
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-4323

Query Match 78.3%; Score 1245; DB 8; Length 2530;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1495; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
```

```
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514


RESULT 13
US-10-843-641A-7148
; Sequence 7148, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7148
;  LENGTH: 2530
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (1)..(2530)
;  OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-7148

  Query Match             78.3%;  Score 1245;  DB 9;  Length 2530;
  Best Local Similarity   99.7%;  Pred. No. 0;

  Matches 1495;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094
```

```
Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1215 AGTTCTAGCATTTCCAGCAGCTCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514
```

RESULT 14
US-10-883-805-12
; Sequence 12, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: LAL, Preeti
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
; FILE REFERENCE: PF-0567 PCT
; CURRENT APPLICATION NUMBER: US/10/883,805
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 09/123,531
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 09/206,817
; PRIOR FILING DATE: 1998-12-07
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 2284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-12

Query Match 61.5%; Score 978; DB 9; Length 2284;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db         47 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGTTGGCACTG 106

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 166

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 226

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 286

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 346

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 406

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 466

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 526

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 586

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAG 646

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACCTTTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 706

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        707 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCCGGCCCCTACATC 766

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 826

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 886

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 946

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        947 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTACCCT 1006

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1066

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1126

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1127 TCGGCAGCTTCCTCGGCCATTGCATTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1186
```

```
Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1246

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
             |||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db       1247 AGTTCTAGCATTTCCAGCAGCTCCGGTTTACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1306

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1307 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCCAGT 1366

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1367 GGCAAAATCATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1426

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1486

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1546

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1606

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
             ||||||||||||||||||||||||||||||
Db       1607 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1636
```

```
RESULT 15
US-10-164-230-60
; Sequence 60, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
;   LENGTH: 1725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (15)...(1601)
US-10-164-230-60

  Query Match             60.8%;  Score 966;  DB 6;  Length 1725;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1566;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||| |||||||||||||||||||| |||| ||||||||||||||||||||||||||
Db        435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||||| || |||||||||||||||||||||||||||||||||||
Db       1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274
```

```
Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTC 1578
             ||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTC 1592
```

RESULT 16
US-10-723-860-8160
; Sequence 8160, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8160
; LENGTH: 2551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (239)..(313)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (476)..(539)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1564)..(1593)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-8160

Query Match 58.7%; Score 933; DB 8; Length 2551;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 983; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 637

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 697

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 757

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 817

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GCCCCTGGAGTGGTTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 877

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 937

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 997

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1057

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1117

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1118 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1177

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1178 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1237

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1238 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1297

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1298 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1357

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1358 ATCCTTCAGCCTTGCGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1417

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1418 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1477

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1478 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1537

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCC 1533
              |||||||||||||||||||||||||
Db       1538 GCCCAAGTGAAGCCTCTGGGGCCC 1561
```

RESULT 17
US-10-204-884-2
; Sequence 2, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS

```
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1590
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1590)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (447)..(450)
;   OTHER INFORMATION: Ser deletion
US-10-204-884-2

  Query Match             39.1%;  Score 621;  DB 6;  Length 1590;
  Best Local Similarity   98.9%;  Pred. No. 1.9e-298;
  Matches 1371;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACT 372
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGT 372

Qy        373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGC 432
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        373 CTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAARCAGC 432

Qy        433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGC 492

Qy        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGC 552
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        493 AGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGC 552

Qy        553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTA 612
              |||||||||||||||||||||||||||||||||||||||||||||||| |||| |||||||
Db        553 GGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCARACCTYTGGGGTA 612

Qy        613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 TCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGAC 672

Qy        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCYGGCCCCTACATCCCCAGCTCCCAC 732

Qy        733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        733 TCTGTGTCAGGGGGTCAGAGRCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792

Qy        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852

Qy        853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGAC 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGAC 912

Qy        913 AGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTC 972
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        913 AGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTAYCCTGTGGGCTACTTC 972

Qy        973 ACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCC 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 ACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCC 1032

Qy       1033 ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCC 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 ATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCC 1092

Qy       1093 TCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCC 1152
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 TCGGCCATTGCRTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCC 1152

Qy       1153 ACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATT 1212
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1153 ACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCRGTTCTAGCATT 1212

Qy       1213 TCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCC 1272
              ||||||||| ||| || ||||||||||||||||||||||||||||||||||||||||||
Db       1213 TCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCC 1272

Qy       1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATC 1332
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1273 CCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCSAGTGGCAAAATCATC 1332

Qy       1333 CTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCC 1392
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 CTTCAGCCTTGYGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCC 1392

Qy       1393 TTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAG 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1393 TTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAG 1452

Qy       1453 CCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCC 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1453 CCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCC 1512

Qy       1513 CAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAG 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1513 CAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAG 1572

Qy       1573 TTACTC 1578
              ||||||
Db       1573 TTACTC 1578


RESULT 18
US-10-029-386-25367
; Sequence 25367, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25367
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000510.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.21
;   OTHER INFORMATION: NT HIT: gi11419124, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BG676174.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q15517, EVALUE 7.00e-25
```

```
US-10-029-386-25367

  Query Match             35.0%;  Score 557;  DB 6;  Length 659;
  Best Local Similarity   99.7%;  Pred. No. 1.5e-266;
  Matches  657;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          86 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 145
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCA 60

Qy         146 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACA 120

Qy         206 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 265
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCA 180

Qy         266 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAG 325
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAG 240

Qy         326 GAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTGCAT 385
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         241 GAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTGCAT 300

Qy         386 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGG 445
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGG 360

Qy         446 GAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 505
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCTTCC 420

Qy         506 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAAACC 565
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         421 AAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAAACC 480

Qy         566 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 625
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTGGCC 540

Qy         626 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 685
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 AAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTG 600

Qy         686 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 744
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGG 659


RESULT 19
US-10-029-386-11663
; Sequence 11663, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11663
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000510.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.21
;   OTHER INFORMATION: EST_HUMAN HIT: BF834417.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: Q15517, EVALUE 8.00e-59
;   OTHER INFORMATION: NT HIT: AF224753.1, EVALUE 0.00e+00
US-10-029-386-11663

  Query Match             31.8%;  Score 505;  DB 6;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-240;
  Matches  505;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         629 GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAG 688
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCAGTGGAG 60

Qy         689 GGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTC 748
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTC 120

Qy         749 AGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAG 808
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGGTTCAAG 180

Qy         809 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 868
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAG 240

Qy         869 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 928
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAG 300

Qy         929 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 988
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGAACCCTG 360

Qy         989 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 1048
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAAT 420

Qy        1049 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 1108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGCGTTCC 480

Qy        1109 AGCCAGTGGGGACTGGTGGGGTCCA 1133
               |||||||||||||||||||||||||
Db         481 AGCCAGTGGGGACTGGTGGGGTCCA 505


RESULT 20
US-09-925-065A-244963
; Sequence 244963, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244963
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-09-925-065A-244963

  Query Match             18.1%;  Score 288;  DB 4;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-132;
  Matches  288;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 375

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 370
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db         496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 543


RESULT 21
US-09-925-065A-244955
; Sequence 244955, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244955
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244955

  Query Match             15.5%;  Score 247;  DB 4;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-112;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         297 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 356

Qy         184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         357 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 416

Qy         244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         417 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 476

Qy         304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         477 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 536

Qy         364 GGCTCTA 370
               |||||||
Db         537 GGCTCTA 543


RESULT 22
US-09-925-065A-244956
; Sequence 244956, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244956
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244956

  Query Match             14.9%;  Score 237;  DB 4;  Length 580;
  Best Local Similarity   99.7%;  Pred. No. 3.3e-107;
  Matches  287;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGYTTCAGTAGCT 375

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 370
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db         496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 543


RESULT 23
US-10-316-175-67
; Sequence 67, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
```

```
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 67
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-316-175-67

  Query Match             5.3%;  Score 85;  DB 6;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 2e-31;
  Matches   85;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAG 85
              |||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAG 99


RESULT 24
US-10-204-884-145
; Sequence 145, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 145
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-145

  Query Match             5.0%;  Score 79;  DB 6;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 2e-28;
  Matches   79;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 GGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATT 463
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATT 60

Qy        464 CGAGCAGCAGCAGCAGCTT 482
              |||||||||||||||||||
Db         61 CGAGCAGCAGCAGCAGCTT 79


RESULT 25
US-10-204-884-26
; Sequence 26, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-26

  Query Match             3.1%;  Score 50;  DB 6;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-14;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        425 GAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGC 50


RESULT 26
US-10-204-884-147
; Sequence 147, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 85
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-147

  Query Match             3.1%;  Score 50;  DB 6;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-14;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCT 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCT 50


RESULT 27
US-10-204-884-150
; Sequence 150, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
```

```
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 150
;   LENGTH: 81
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-150

  Query Match             2.9%;  Score 46;  DB 6;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-12;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        911 ACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTA 956
              ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTA 46


RESULT 28
US-10-204-884-146
; Sequence 146, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
;   LENGTH: 85
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-146

  Query Match             2.8%;  Score 45;  DB 6;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-11;
  Matches   45;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        555 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 599
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCA 45


RESULT 29
US-10-204-884-151
; Sequence 151, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-151

  Query Match             2.7%;  Score 43;  DB 6;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1061 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 1103
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 43


RESULT 30
US-10-467-752-74
; Sequence 74, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-74

  Query Match             2.7%;  Score 43;  DB 7;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-10;
  Matches   43;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1061 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 1103
              |||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTGC 43


RESULT 31
US-10-204-884-144
; Sequence 144, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 144
;   LENGTH: 80
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-144

  Query Match             2.6%;  Score 42;  DB 6;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-10;
  Matches   42;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        386 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAA 427
              ||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAA 42
```

```
RESULT 32
US-10-204-884-152
; Sequence 152, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 152
;   LENGTH: 76
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-152

  Query Match             2.6%;  Score 41;  DB 6;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-09;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              |||||||||||||||||||||||||||||||||||||||||
Db          1 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 41


RESULT 33
US-10-467-752-73
; Sequence 73, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
;   LENGTH: 76
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-73

  Query Match             2.6%;  Score 41;  DB 7;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-09;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              |||||||||||||||||||||||||||||||||||||||||
Db          1 CCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 41


RESULT 34
US-10-204-884-143
; Sequence 143, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-143

  Query Match             2.6%;  Score 41;  DB 6;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-09;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 233
              |||||||||||||||||||||||||||||||||||||||||
Db         39 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCAT 79


RESULT 35
US-10-467-752-76
; Sequence 76, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
;   LENGTH: 81
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-76

  Query Match             2.6%;  Score 41;  DB 7;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-09;
  Matches   41;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        123 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCT 163
              |||||||||||||||||||||||||||||||||||||||||
Db         41 CTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCT 81


RESULT 36
US-10-204-884-149
; Sequence 149, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 149
; LENGTH: 79
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-149

Query Match 2.5%; Score 40; DB 6; Length 79;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 713 CCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAG 752
||||||||||||||||||||||||||||||||||||||||
Db 1 CCTACATCCCCAGCTCCCACTCTGTGTCAGGGGGTCAGAG 40

RESULT 37
US-10-204-884-141
; Sequence 141, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 81
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-141

Query Match 2.5%; Score 40; DB 6; Length 81;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 122
||||||||||||||||||||||||||||||||||||||||
Db 1 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCC 40

RESULT 38
US-10-204-884-148
; Sequence 148, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-148

Query Match 2.5%; Score 39; DB 6; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 CGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTC 707
|||||||||||||||||||||||||||||||||||||||
Db 1 CGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTC 39

RESULT 39
US-10-467-752-75
; Sequence 75, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
; APPLICANT: Lench, et al.
; TITLE OF INVENTION: Test and Model for Inflammatory disease
; FILE REFERENCE: 2003882-0009
; CURRENT APPLICATION NUMBER: US/10/467,752
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: PCT/GB02/00653
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: GB0103514.6
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 71
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-467-752-75

Query Match 2.4%; Score 38; DB 7; Length 71;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 958 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGG 995
|||||||||||||||||||||||||||||||||||||
Db 34 CCTGTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGG 71

RESULT 40
US-10-204-884-157
; Sequence 157, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-204-884-157

Query Match 2.4%; Score 38; DB 6; Length 72;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1306 AGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTG 1343
|||||||||||||||||||||||||||||||||||||
Db 1 AGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTG 38

```
RESULT 41
US-10-204-884-156
; Sequence 156, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 156
;   LENGTH: 74
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-156

  Query Match             2.4%;  Score 38;  DB 6;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-08;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 1316
             |||||||||||||||||||||||||||||||||||||||
Db          1 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 38


RESULT 42
US-10-467-752-70
; Sequence 70, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
;   LENGTH: 74
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-70

  Query Match             2.4%;  Score 38;  DB 7;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-08;
  Matches   38;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1279 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 1316
             |||||||||||||||||||||||||||||||||||||||
Db          1 GGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATC 38


RESULT 43
US-10-204-884-142
; Sequence 142, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-142

  Query Match             2.3%;  Score 37;  DB 6;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-07;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        129 GGACCCCACGCGTATCACCTCCCCTAACGACCCCTGC 165
              |||||||||||||||||||||||||||||||||||||
Db          1 GGACCCCACGCGTATCACCTCCCCTAACGACCCCTGC 37


RESULT 44
US-10-204-884-127
; Sequence 127, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 127
;   LENGTH: 71
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-127

  Query Match             2.3%;  Score 36;  DB 6;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-07;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 GCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATG 51
              ||||||||||||||||||||||||||||||||||||
Db          1 GCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATG 36


RESULT 45
US-10-204-884-153
; Sequence 153, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
```

```
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 153
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-153

  Query Match             2.2%;  Score 35;  DB 6;  Length 69;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-06;
  Matches   35;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1187 CTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGC 1221
              |||||||||||||||||||||||||||||||||||
Db          1 CTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGC 35


RESULT 46
US-10-204-884-155
; Sequence 155, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 155
;   LENGTH: 67
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-155

  Query Match             2.1%;  Score 33;  DB 6;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1230 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 1262
              |||||||||||||||||||||||||||||||||
Db         35 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 67


RESULT 47
US-10-467-752-71
; Sequence 71, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
;   LENGTH: 67
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-71

  Query Match             2.1%;  Score 33;  DB 7;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1230 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 1262
              |||||||||||||||||||||||||||||||||
Db         35 ACCCTACCATCCCTGCGGCAGTGCTTCCCAGAG 67


RESULT 48
US-10-204-884-126
; Sequence 126, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 126
;   LENGTH: 68
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-204-884-126

  Query Match             2.1%;  Score 33;  DB 6;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-05;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGT 33
              |||||||||||||||||||||||||||||||||
Db         36 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGT 68


RESULT 49
US-10-467-752-69
; Sequence 69, Application US/10467752
; Publication No. US20040161759A1
; GENERAL INFORMATION:
;  APPLICANT: Lench, et al.
;  TITLE OF INVENTION: Test and Model for Inflammatory disease
;  FILE REFERENCE: 2003882-0009
;  CURRENT APPLICATION NUMBER: US/10/467,752
;  CURRENT FILING DATE:  2003-08-13
;  PRIOR APPLICATION NUMBER: PCT/GB02/00653
;  PRIOR FILING DATE: 2002-02-13
;  PRIOR APPLICATION NUMBER: GB0103514.6
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
;   LENGTH: 72
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-467-752-69

  Query Match             2.1%;  Score 33;  DB 7;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1345 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 1377
              |||||||||||||||||||||||||||||||||
Db         40 GGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGC 72


RESULT 50
```

US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 2.1%; Score 33; DB 6; Length 3673778;
Best Local Similarity 100.0%; Pred. No. 3.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        131 ACCCCACGCGTATCACCTCCCCTAACGACCCCT 163
              |||||||||||||||||||||||||||||||||
Db    1405203 ACCCCACGCGTATCACCTCCCCTAACGACCCCT 1405171
```

Search completed: March 22, 2006, 03:29:50
Job time : 1378 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 03:06:55 ; Search time 580 Seconds
(without alignments)
6392.090 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 8023312 seqs, 1165852854 residues

Word size : 0

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	288	18.1	580	6	US-09-925-065A-244963	Sequence 244963,
	2	247	15.5	580	6	US-09-925-065A-244955	Sequence 244955,
	3	237	14.9	580	6	US-09-925-065A-244956	Sequence 244956,
	4	25	1.6	25	12	US-11-121-849-122635	Sequence 122635,
	5	25	1.6	25	12	US-11-121-849-122636	Sequence 122636,
	6	25	1.6	25	12	US-11-121-849-122637	Sequence 122637,
	7	25	1.6	25	12	US-11-121-849-122638	Sequence 122638,
	8	25	1.6	25	12	US-11-121-849-122639	Sequence 122639,
	9	25	1.6	25	12	US-11-121-849-122640	Sequence 122640,
c	10	24	1.5	24	12	US-11-069-908-1447	Sequence 1447, Ap
	11	24	1.5	24	12	US-11-069-908-3615	Sequence 3615, Ap
c	12	24	1.5	24	12	US-11-069-908-6179	Sequence 6179, Ap
	13	24	1.5	1475	8	US-10-750-185-59559	Sequence 59559, A
	14	24	1.5	1475	8	US-10-750-623-59559	Sequence 59559, A
	15	23	1.4	24	12	US-11-069-908-1249	Sequence 1249, Ap
c	16	23	1.4	24	12	US-11-069-908-3813	Sequence 3813, Ap
	17	22	1.4	22	12	US-11-069-908-1723	Sequence 1723, Ap
c	18	22	1.4	22	12	US-11-069-908-3143	Sequence 3143, Ap
c	19	21	1.3	22	12	US-11-069-908-777	Sequence 777, App
	20	21	1.3	22	12	US-11-069-908-4089	Sequence 4089, Ap
c	21	21	1.3	24	8	US-10-310-914A-1097919	Sequence 1097919,
	22	21	1.3	544	6	US-09-925-065A-664590	Sequence 664590,
	23	21	1.3	2711	8	US-10-750-185-62762	Sequence 62762, A
	24	21	1.3	2711	8	US-10-750-623-62762	Sequence 62762, A
	25	21	1.3	67858	8	US-10-995-561-13484	Sequence 13484, A
	26	20	1.3	20	12	US-11-069-908-1130	Sequence 1130, Ap
	27	20	1.3	20	12	US-11-069-908-2961	Sequence 2961, Ap
c	28	20	1.3	20	12	US-11-069-908-3140	Sequence 3140, Ap
	29	20	1.3	20	12	US-11-069-908-5327	Sequence 5327, Ap
	30	20	1.3	20	12	US-11-069-908-6455	Sequence 6455, Ap
c	31	20	1.3	474	9	US-11-116-881A-2076	Sequence 2076, Ap
c	32	20	1.3	545	6	US-09-925-065A-588699	Sequence 588699,
	33	20	1.3	838	8	US-10-750-185-26693	Sequence 26693, A
	34	20	1.3	838	8	US-10-750-623-26693	Sequence 26693, A
c	35	20	1.3	2475	8	US-10-750-185-56310	Sequence 56310, A
c	36	20	1.3	2475	8	US-10-750-623-56310	Sequence 56310, A
c	37	20	1.3	3054	8	US-10-750-185-51216	Sequence 51216, A
c	38	20	1.3	3054	8	US-10-750-623-51216	Sequence 51216, A
c	39	20	1.3	4841	8	US-10-750-185-53007	Sequence 53007, A
c	40	20	1.3	4841	8	US-10-750-623-53007	Sequence 53007, A
	41	19	1.2	19	10	US-11-101-244-91681	Sequence 91681, A
	42	19	1.2	19	10	US-11-101-244-91682	Sequence 91682, A
	43	19	1.2	19	10	US-11-101-244-91683	Sequence 91683, A
	44	19	1.2	19	10	US-11-101-244-91685	Sequence 91685, A
	45	19	1.2	19	10	US-11-101-244-91686	Sequence 91686, A
	46	19	1.2	19	10	US-11-101-244-91689	Sequence 91689, A
	47	19	1.2	19	10	US-11-101-244-91691	Sequence 91691, A
	48	19	1.2	19	10	US-11-101-244-91692	Sequence 91692, A
	49	19	1.2	19	10	US-11-101-244-91694	Sequence 91694, A
	50	19	1.2	19	10	US-11-101-244-91699	Sequence 91699, A
	51	19	1.2	19	10	US-11-101-244-91700	Sequence 91700, A
	52	19	1.2	19	10	US-11-101-244-91703	Sequence 91703, A
	53	19	1.2	19	10	US-11-101-244-91704	Sequence 91704, A
	54	19	1.2	19	10	US-11-101-244-91705	Sequence 91705, A
	55	19	1.2	19	10	US-11-101-244-91707	Sequence 91707, A
	56	19	1.2	19	10	US-11-101-244-91711	Sequence 91711, A
	57	19	1.2	19	10	US-11-101-244-91712	Sequence 91712, A
	58	19	1.2	19	10	US-11-101-244-91713	Sequence 91713, A
	59	19	1.2	19	10	US-11-101-244-91717	Sequence 91717, A
	60	19	1.2	19	10	US-11-101-244-91718	Sequence 91718, A
	61	19	1.2	19	10	US-11-101-244-91719	Sequence 91719, A
	62	19	1.2	19	10	US-11-101-244-91720	Sequence 91720, A
	63	19	1.2	19	10	US-11-101-244-91725	Sequence 91725, A
	64	19	1.2	19	10	US-11-101-244-91726	Sequence 91726, A
	65	19	1.2	19	10	US-11-101-244-91728	Sequence 91728, A
	66	19	1.2	19	10	US-11-101-244-91729	Sequence 91729, A
	67	19	1.2	19	10	US-11-101-244-91730	Sequence 91730, A
	68	19	1.2	19	10	US-11-101-244-91731	Sequence 91731, A
	69	19	1.2	19	10	US-11-101-244-91732	Sequence 91732, A
	70	19	1.2	19	10	US-11-101-244-91733	Sequence 91733, A
	71	19	1.2	19	10	US-11-101-244-91734	Sequence 91734, A
	72	19	1.2	19	10	US-11-101-244-91735	Sequence 91735, A
	73	19	1.2	19	10	US-11-101-244-91738	Sequence 91738, A
	74	19	1.2	19	10	US-11-101-244-91739	Sequence 91739, A
	75	19	1.2	19	10	US-11-101-244-91743	Sequence 91743, A

ALIGNMENTS

RESULT 1
US-09-925-065A-244963
; Sequence 244963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244963
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244963

  Query Match             18.1%;  Score 288;  DB 6;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-138;
  Matches  288;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 375

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 370
               |||||||||||||||||||||||||||||||||||||||||||||||
Db         496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 543


RESULT 2
US-09-925-065A-244955
; Sequence 244955, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244955
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244955

  Query Match             15.5%;  Score 247;  DB 6;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-117;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         124 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         297 TGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGAC 356

Qy         184 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         357 TCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCC 416

Qy         244 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         417 AGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGT 476

Qy         304 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         477 TCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCT 536

Qy         364 GGCTCTA 370
               |||||||
Db         537 GGCTCTA 543


RESULT 3
US-09-925-065A-244956
; Sequence 244956, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244956
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-244956

  Query Match             14.9%;  Score 237;  DB 6;  Length 580;
  Best Local Similarity   99.7%;  Pred. No. 8.7e-112;
  Matches  287;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         256 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 315

Qy         143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
               |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         316 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGYTTCAGTAGCT 375

Qy         203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         376 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 435

Qy         263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         436 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 495

Qy         323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 370
               |||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        496 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTA 543


RESULT 4
US-11-121-849-122635
; Sequence 122635, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122635
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122635

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1383 TGTCTCCTCCTTGACACTGACTGGG 1407
             |||||||||||||||||||||||||
Db          1 TGTCTCCTCCTTGACACTGACTGGG 25


RESULT 5
US-11-121-849-122636
; Sequence 122636, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122636
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122636

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1438 TCCGCTGGTGCCAAGCCCTGTGGCT 1462
             |||||||||||||||||||||||||
Db          1 TCCGCTGGTGCCAAGCCCTGTGGCT 25


RESULT 6
US-11-121-849-122637
; Sequence 122637, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122637
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122637

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1454 CCTGTGGCTCCAGCAGTGCTGGAAA 1478
             |||||||||||||||||||||||||
Db          1 CCTGTGGCTCCAGCAGTGCTGGAAA 25


RESULT 7
US-11-121-849-122638
; Sequence 122638, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122638
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122638

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1529 GGCCCCAGCTAGCTGACCCTGAAGT 1553
             |||||||||||||||||||||||||
Db          1 GGCCCCAGCTAGCTGACCCTGAAGT 25


RESULT 8
US-11-121-849-122639
; Sequence 122639, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122639
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122639

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
```

```
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1548 TGAAGTTTTCCTACCCCAAGGAGAG 1572
              |||||||||||||||||||||||||
Db          1 TGAAGTTTTCCTACCCCAAGGAGAG 25


RESULT 9
US-11-121-849-122640
; Sequence 122640, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 122640
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-122640

  Query Match             1.6%;  Score 25;  DB 12;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1558 CTACCCCAAGGAGAGTTACTCGACA 1582
              |||||||||||||||||||||||||
Db          1 CTACCCCAAGGAGAGTTACTCGACA 25


RESULT 10
US-11-069-908-1447/c
; Sequence 1447, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1447
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-1447

  Query Match             1.5%;  Score 24;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.088;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        957 TCCTGTGGGCTACTTCACCAAAGA 980
              ||||||||||||||||||||||||
Db         24 TCCTGTGGGCTACTTCACCAAAGA 1


RESULT 11
US-11-069-908-3615
; Sequence 3615, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3615
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-3615

  Query Match             1.5%;  Score 24;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.088;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        730 CACTCTGTGTCAGGGGGTCAGAGG 753
              ||||||||||||||||||||||||
Db          1 CACTCTGTGTCAGGGGGTCAGAGG 24


RESULT 12
US-11-069-908-6179/c
; Sequence 6179, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6179
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-6179

  Query Match             1.5%;  Score 24;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.088;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        932 TGACCTACAGTAAGGGTAAAATCT 955
              ||||||||||||||||||||||||
Db         24 TGACCTACAGTAAGGGTAAAATCT 1


RESULT 13
US-10-750-185-59559
; Sequence 59559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
```

```
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 59559
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Bovine  19866881051537
US-10-750-185-59559

  Query Match             1.5%;  Score 24;  DB 8;  Length 1475;
  Best Local Similarity   100.0%;  Pred. No. 0.096;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         59 TGCTGCTGGCTGGTCTCCTCCTGC 82
              ||||||||||||||||||||||||
Db        279 TGCTGCTGGCTGGTCTCCTCCTGC 302


RESULT 14
US-10-750-623-59559
; Sequence 59559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 59559
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Bovine  19866881051537
US-10-750-623-59559

  Query Match             1.5%;  Score 24;  DB 8;  Length 1475;
  Best Local Similarity   100.0%;  Pred. No. 0.096;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         59 TGCTGCTGGCTGGTCTCCTCCTGC 82
              ||||||||||||||||||||||||
Db        279 TGCTGCTGGCTGGTCTCCTCCTGC 302


RESULT 15
US-11-069-908-1249
; Sequence 1249, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1249
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-1249

  Query Match             1.4%;  Score 23;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        730 CACTCTGTGTCAGGGGGTCAGAG 752
              |||||||||||||||||||||||
Db          1 CACTCTGTGTCAGGGGGTCAGAG 23


RESULT 16
US-11-069-908-3813/c
; Sequence 3813, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3813
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-3813

  Query Match             1.4%;  Score 23;  DB 12;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        958 CCTGTGGGCTACTTCACCAAAGA 980
              |||||||||||||||||||||||
Db         23 CCTGTGGGCTACTTCACCAAAGA 1


RESULT 17
US-11-069-908-1723
; Sequence 1723, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1723
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-1723

  Query Match             1.4%;  Score 22;  DB 12;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.95;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1296 CAGCAGCTCCAGTTCCCAATCG 1317
             ||||||||||||||||||||||
Db          1 CAGCAGCTCCAGTTCCCAATCG 22


RESULT 18
US-11-069-908-3143/c
; Sequence 3143, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3143
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-3143

  Query Match             1.4%;  Score 22;  DB 12;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.95;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        123 CTGTAAGGACCCCACGCGTATC 144
             ||||||||||||||||||||||
Db         22 CTGTAAGGACCCCACGCGTATC 1


RESULT 19
US-11-069-908-777/c
; Sequence 777, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 777
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-777

  Query Match             1.3%;  Score 21;  DB 12;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        124 TGTAAGGACCCCACGCGTATC 144
             |||||||||||||||||||||
Db         21 TGTAAGGACCCCACGCGTATC 1


RESULT 20
US-11-069-908-4089
; Sequence 4089, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4089
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-4089

  Query Match             1.3%;  Score 21;  DB 12;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1296 CAGCAGCTCCAGTTCCCAATC 1316
             |||||||||||||||||||||
Db          1 CAGCAGCTCCAGTTCCCAATC 21


RESULT 21
US-10-310-914A-1097919/c
; Sequence 1097919, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
;  APPLICANT: Bentwich, Isaac
;  APPLICANT:  Shiler, Kvuzat
;  TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
;  TITLE OF INVENTION:  uses thereof
;  FILE REFERENCE: 06087.0200.CPUS01
;  CURRENT APPLICATION NUMBER: US/10/310,914A
;  CURRENT FILING DATE:  2002-12-06
;  NUMBER OF SEQ ID NOS: 1388402
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1097919
;   LENGTH: 24
;   TYPE: RNA
;   ORGANISM: Human
US-10-310-914A-1097919

  Query Match             1.3%;  Score 21;  DB 8;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 3.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1285 GGCTCCTTCAGCAGCAGCTCC 1305
             |||||||||||||||||||||
Db         22 GGCTCCTTCAGCAGCAGCTCC 2


RESULT 22
US-09-925-065A-664590
; Sequence 664590, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 664590
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-664590

Query Match 1.3%; Score 21; DB 6; Length 544;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 752 GGCCTGTGGTGGTGGTGGTGG 772
||||||||||||||||||||
Db 36 GGCCTGTGGTGGTGGTGGTGG 56

RESULT 23
US-10-750-185-62762
; Sequence 62762, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 62762
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Bovine 19866880740241
US-10-750-185-62762

Query Match 1.3%; Score 21; DB 8; Length 2711;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 527 TGCCAACCAATGACAACTCTT 547
||||||||||||||||||||
Db 137 TGCCAACCAATGACAACTCTT 157

RESULT 24
US-10-750-623-62762
; Sequence 62762, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 62762
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Bovine 19866880740241
US-10-750-623-62762

Query Match 1.3%; Score 21; DB 8; Length 2711;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 527 TGCCAACCAATGACAACTCTT 547
||||||||||||||||||||
Db 137 TGCCAACCAATGACAACTCTT 157

RESULT 25
US-10-995-561-13484
; Sequence 13484, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13484
; LENGTH: 67858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(67858)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13484

Query Match 1.3%; Score 21; DB 8; Length 67858;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1285 GGCTCCTTCAGCAGCAGCTCC 1305
||||||||||||||||||||
Db 64590 GGCTCCTTCAGCAGCAGCTCC 64610

RESULT 26
US-11-069-908-1130
; Sequence 1130, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
; APPLICANT: OLIPHANT, ARNOLD
; APPLICANT: MURRAY, SARAH
; TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
; TITLE OF INVENTION: CONDITIONS
; FILE REFERENCE: 029011-0402

```
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1130
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-1130

  Query Match             1.3%;  Score 20;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        689 GGCCCATCGTCTCGCACTCT 708
              ||||||||||||||||||||
Db          1 GGCCCATCGTCTCGCACTCT 20


RESULT 27
US-11-069-908-2961
; Sequence 2961, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2961
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-2961

  Query Match             1.3%;  Score 20;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        173 GGAAGGGTGACTCCAGCGGC 192
              ||||||||||||||||||||
Db          1 GGAAGGGTGACTCCAGCGGC 20


RESULT 28
US-11-069-908-3140/c
; Sequence 3140, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3140
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-3140

  Query Match             1.3%;  Score 20;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        605 CTGGGGTATCCAGCAGTGGC 624
              ||||||||||||||||||||
Db         20 CTGGGGTATCCAGCAGTGGC 1


RESULT 29
US-11-069-908-5327
; Sequence 5327, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5327
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-5327

  Query Match             1.3%;  Score 20;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        194 TCAGTAGCTACAGTGGCTCC 213
              ||||||||||||||||||||
Db          1 TCAGTAGCTACAGTGGCTCC 20


RESULT 30
US-11-069-908-6455
; Sequence 6455, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
;  APPLICANT: OLIPHANT, ARNOLD
;  APPLICANT:  MURRAY, SARAH
;  TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
;  TITLE OF INVENTION:  CONDITIONS
;  FILE REFERENCE: 029011-0402
;  CURRENT APPLICATION NUMBER: US/11/069,908
;  CURRENT FILING DATE:  2005-02-28
;  PRIOR APPLICATION NUMBER: 60/547,823
;  PRIOR FILING DATE: 2004-02-26
;  NUMBER OF SEQ ID NOS: 7098
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6455
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-6455
```

Query Match 1.3%; Score 20; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1319 GTGGCAAAATCATCCTTCAG 1338
              ||||||||||||||||||||
Db          1 GTGGCAAAATCATCCTTCAG 20
```

RESULT 31
US-11-116-881A-2076/c
; Sequence 2076, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2076
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
US-11-116-881A-2076

Query Match 1.3%; Score 20; DB 9; Length 474;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        467 GCAGCAGCAGCAGCTTTCAG 486
              ||||||||||||||||||||
Db        467 GCAGCAGCAGCAGCTTTCAG 448
```

RESULT 32
US-09-925-065A-588699/c
; Sequence 588699, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 588699
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-588699

Query Match 1.3%; Score 20; DB 6; Length 545;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        221 CTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||
Db        164 CTGGCAGCTCCATTTCCAGT 145
```

RESULT 33
US-10-750-185-26693
; Sequence 26693, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26693
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Bovine 19866881820487
US-10-750-185-26693

Query Match 1.3%; Score 20; DB 8; Length 838;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1293 CAGCAGCAGCTCCAGTTCCC 1312
              ||||||||||||||||||||
Db        294 CAGCAGCAGCTCCAGTTCCC 313
```

RESULT 34
US-10-750-623-26693
; Sequence 26693, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623

```
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26693
;   LENGTH: 838
;   TYPE: DNA
;   ORGANISM: Bovine  19866881820487
US-10-750-623-26693

  Query Match             1.3%;  Score 20;  DB 8;  Length 838;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1293 CAGCAGCAGCTCCAGTTCCC 1312
              ||||||||||||||||||||
Db        294 CAGCAGCAGCTCCAGTTCCC 313


RESULT 35
US-10-750-185-56310/c
; Sequence 56310, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 56310
;   LENGTH: 2475
;   TYPE: DNA
;   ORGANISM: Bovine  19866880425065
US-10-750-185-56310

  Query Match             1.3%;  Score 20;  DB 8;  Length 2475;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        466 AGCAGCAGCAGCAGCTTTCA 485
              ||||||||||||||||||||
Db       2280 AGCAGCAGCAGCAGCTTTCA 2261


RESULT 36
US-10-750-623-56310/c
; Sequence 56310, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 56310
;   LENGTH: 2475
;   TYPE: DNA
;   ORGANISM: Bovine  19866880425065
US-10-750-623-56310

  Query Match             1.3%;  Score 20;  DB 8;  Length 2475;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        466 AGCAGCAGCAGCAGCTTTCA 485
              ||||||||||||||||||||
Db       2280 AGCAGCAGCAGCAGCTTTCA 2261


RESULT 37
US-10-750-185-51216/c
; Sequence 51216, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 51216
;   LENGTH: 3054
;   TYPE: DNA
;   ORGANISM: Bovine  19866880413398
US-10-750-185-51216

  Query Match             1.3%;  Score 20;  DB 8;  Length 3054;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        466 AGCAGCAGCAGCAGCTTTCA 485
              ||||||||||||||||||||
Db       2633 AGCAGCAGCAGCAGCTTTCA 2614


RESULT 38
US-10-750-623-51216/c
; Sequence 51216, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 51216
```

; LENGTH: 3054
; TYPE: DNA
; ORGANISM: Bovine 19866880413398
US-10-750-623-51216

Query Match 1.3%; Score 20; DB 8; Length 3054;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 466 AGCAGCAGCAGCAGCTTTCA 485
||||||||||||||||||||
Db 2633 AGCAGCAGCAGCAGCTTTCA 2614

RESULT 39
US-10-750-185-53007/c
; Sequence 53007, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53007
; LENGTH: 4841
; TYPE: DNA
; ORGANISM: Bovine 19866880539163
US-10-750-185-53007

Query Match 1.3%; Score 20; DB 8; Length 4841;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 466 AGCAGCAGCAGCAGCTTTCA 485
||||||||||||||||||||
Db 2574 AGCAGCAGCAGCAGCTTTCA 2555

RESULT 40
US-10-750-623-53007/c
; Sequence 53007, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53007
; LENGTH: 4841
; TYPE: DNA
; ORGANISM: Bovine 19866880539163
US-10-750-623-53007

Query Match 1.3%; Score 20; DB 8; Length 4841;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 466 AGCAGCAGCAGCAGCTTTCA 485
||||||||||||||||||||
Db 2574 AGCAGCAGCAGCAGCTTTCA 2555

RESULT 41
US-11-101-244-91681
; Sequence 91681, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91681
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91681

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 68.4%; Pred. No. 34;
Matches 13; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 577 GGACAAAGCTCTTCCTCTT 595
||||||||:|::||:|::
Db 1 GGACAAAGCUCUUCCUCUU 19

RESULT 42
US-11-101-244-91682
; Sequence 91682, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91682
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91682

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 78.9%; Pred. No. 34;
Matches 15; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1043 GCAAATACTTCTCCAGCAA 1061
              |||||:||::|:|||||||
Db          1 GCAAAUACUUCUCCAGCAA 19
```

RESULT 43
US-11-101-244-91683
; Sequence 91683, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91683
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91683

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 73.7%; Pred. No. 34;
Matches 14; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy        253 GGTGGTGGCTCCAGTGGTA 271
              ||:||:|||:||||:||:|
Db          1 GGUGGUGGCUCCAGUGGUA 19
```

RESULT 44
US-11-101-244-91685
; Sequence 91685, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91685
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91685

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 78.9%; Pred. No. 34;
Matches 15; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy        530 CAACCAATGACAACTCTTA 548
              |||||||:||||||:|::|
Db          1 CAACCAAUGACAACUCUUA 19
```

RESULT 45
US-11-101-244-91686
; Sequence 91686, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91686
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91686

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 78.9%; Pred. No. 34;
Matches 15; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy        445 GGAAGCAGCAGCTCTCATT 463
              |||||||||||||:|:||::
Db          1 GGAAGCAGCAGCUCUCAUU 19
```

RESULT 46
US-11-101-244-91689
; Sequence 91689, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91689
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91689

Query Match 1.2%; Score 19; DB 10; Length 19;

Best Local Similarity 84.2%; Pred. No. 34;
Matches 16; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy         82 CCAGGGACCTTGGCTAAGA 100
              |||||||||::|||:||||
Db          1 CCAGGGACCUUGGCUAAGA 19
```

RESULT 47
US-11-101-244-91691
; Sequence 91691, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91691
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91691

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 89.5%; Pred. No. 34;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy        975 CAAAGAGAACCCTGTGAAA 993
              |||||||||||:|:||||
Db          1 CAAAGAGAACCCUGUGAAA 19
```

RESULT 48
US-11-101-244-91692
; Sequence 91692, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91692
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91692

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 78.9%; Pred. No. 34;
Matches 15; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy        855 AATCACCTCTGTAGACAAA 873
              ||:||||:|:|:|||||||
Db          1 AAUCACCUCUGUAGACAAA 19
```

RESULT 49
US-11-101-244-91694
; Sequence 91694, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91694
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91694

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 78.9%; Pred. No. 34;
Matches 15; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy        518 GCTCTGCTCTGCCAACCAA 536
              ||:|:||:|:|||||||||
Db          1 GCUCUGCUCUGCCAACCAA 19
```

RESULT 50
US-11-101-244-91699
; Sequence 91699, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 91699
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-91699

Query Match 1.2%; Score 19; DB 10; Length 19;
Best Local Similarity 89.5%; Pred. No. 34;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy        974 CCAAAGAGAACCCTGTGAA 992
              ||||||||||||:|:|||
Db          1 CCAAAGAGAACCCUGUGAA 19
```

Search completed: March 22, 2006, 03:39:28
Job time : 581 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2006, 02:57:00 ; Search time 282 Seconds
(without alignments)
10022.422 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 1590
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 75 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1539	96.8	1590	3	US-09-381-656-2	Sequence 2, Appli
2		1355	85.2	5337	3	US-10-164-230-1	Sequence 1, Appli
3		1355	85.2	25235	3	US-10-164-230-2	Sequence 2, Appli
4		966	60.8	1725	3	US-10-164-230-60	Sequence 60, Appl
c 5		22	1.4	22	3	US-10-164-230-15	Sequence 15, Appl
6		21	1.3	21	3	US-10-164-230-14	Sequence 14, Appl
7		21	1.3	222	3	US-09-641-931F-35	Sequence 35, Appl
c 8		21	1.3	384	3	US-09-248-796A-13896	Sequence 13896, A
9		21	1.3	56551	3	US-09-949-016-12030	Sequence 12030, A
c 10		21	1.3	69813	3	US-09-949-016-12455	Sequence 12455, A
c 11		21	1.3	69813	3	US-09-949-016-13905	Sequence 13905, A
c 12		21	1.3	69813	3	US-09-949-016-13906	Sequence 13906, A
c 13		21	1.3	69833	3	US-09-949-016-12861	Sequence 12861, A
c 14		20	1.3	20	3	US-10-164-230-5	Sequence 5, Appli
15		20	1.3	20	3	US-10-164-230-10	Sequence 10, Appl
16		20	1.3	601	3	US-09-949-002-1172	Sequence 1172, Ap
17		20	1.3	601	3	US-09-949-002-7268	Sequence 7268, Ap
c 18		20	1.3	2563	2	US-08-423-441-1	Sequence 1, Appli
19		20	1.3	40325	3	US-09-949-002-580	Sequence 580, App
20		20	1.3	40325	3	US-09-949-002-771	Sequence 771, App
21		19	1.2	19	3	US-10-164-230-12	Sequence 12, Appl
22		19	1.2	50	3	US-09-023-942A-19	Sequence 19, Appl
23		19	1.2	55	2	US-08-348-143-14	Sequence 14, Appl
24		19	1.2	55	2	US-08-571-785-14	Sequence 14, Appl
25		19	1.2	55	3	US-09-192-435-14	Sequence 14, Appl
26		19	1.2	55	3	US-09-558-340-14	Sequence 14, Appl
27		19	1.2	233	3	US-09-270-767-8813	Sequence 8813, Ap
28		19	1.2	233	3	US-09-270-767-24095	Sequence 24095, A
c 29		19	1.2	240	2	US-08-209-747-11	Sequence 11, Appl
c 30		19	1.2	240	2	US-08-458-298-11	Sequence 11, Appl
31		19	1.2	366	3	US-09-854-133-659	Sequence 659, App
32		19	1.2	523	3	US-09-621-976-14545	Sequence 14545, A
c 33		19	1.2	601	3	US-09-949-016-61662	Sequence 61662, A
c 34		19	1.2	601	3	US-09-949-016-61784	Sequence 61784, A
35		19	1.2	601	3	US-09-949-016-94504	Sequence 94504, A
36		19	1.2	601	3	US-09-949-016-94505	Sequence 94505, A
37		19	1.2	601	3	US-09-949-016-94506	Sequence 94506, A
38		19	1.2	601	3	US-09-949-002-1670	Sequence 1670, Ap
39		19	1.2	601	3	US-09-949-002-1671	Sequence 1671, Ap
40		19	1.2	601	3	US-09-949-002-5678	Sequence 5678, Ap
41		19	1.2	601	3	US-09-949-002-5679	Sequence 5679, Ap
c 42		19	1.2	750	3	US-09-248-796A-499	Sequence 499, App
43		19	1.2	751	3	US-09-270-767-5649	Sequence 5649, Ap
44		19	1.2	751	3	US-09-270-767-20931	Sequence 20931, A
45		19	1.2	777	3	US-08-998-416-218	Sequence 218, App
c 46		19	1.2	1843	6	PCT-US95-03323A-1	Sequence 1, Appli
47		19	1.2	2565	3	US-09-799-451-477	Sequence 477, App
c 48		19	1.2	3990	3	US-09-830-230A-629	Sequence 629, App
49		19	1.2	6530	2	US-08-146-930-1	Sequence 1, Appli
50		19	1.2	6530	3	US-08-458-240-1	Sequence 1, Appli
51		19	1.2	6530	6	PCT-US93-03993-1	Sequence 1, Appli
c 52		19	1.2	7766	3	US-09-125-619-3	Sequence 3, Appli
c 53		19	1.2	7766	3	US-10-222-566-3	Sequence 3, Appli
c 54		19	1.2	7766	3	US-10-143-024A-3	Sequence 3, Appli
c 55		19	1.2	7766	3	US-10-222-162-3	Sequence 3, Appli
56		19	1.2	9208	3	US-09-068-506-1	Sequence 1, Appli
c 57		19	1.2	39085	3	US-09-949-016-14479	Sequence 14479, A
c 58		19	1.2	39088	3	US-09-949-016-14373	Sequence 14373, A
59		19	1.2	42000	3	US-10-081-563-25	Sequence 25, Appl
c 60		19	1.2	47751	3	US-09-949-016-12812	Sequence 12812, A
c 61		19	1.2	47752	3	US-09-949-016-16955	Sequence 16955, A
c 62		19	1.2	74545	3	US-09-949-002-606	Sequence 606, App
c 63		19	1.2	87648	3	US-09-949-016-13655	Sequence 13655, A
64		19	1.2	113283	3	US-09-949-016-16976	Sequence 16976, A
65		19	1.2	113283	3	US-09-949-016-16977	Sequence 16977, A
66		19	1.2	119649	3	US-09-949-016-12537	Sequence 12537, A
c 67		19	1.2	131860	3	US-09-949-002-730	Sequence 730, App
c 68		19	1.2	145287	3	US-09-949-016-13530	Sequence 13530, A
c 69		19	1.2	145287	3	US-09-949-016-13531	Sequence 13531, A
70		19	1.2	392000	3	US-10-027-983-11	Sequence 11, Appl
c 71		18	1.1	18	3	US-10-164-230-11	Sequence 11, Appl
c 72		18	1.1	18	3	US-10-164-230-13	Sequence 13, Appl
73		18	1.1	18	3	US-10-164-230-16	Sequence 16, Appl
74		18	1.1	45	3	US-08-721-458B-62	Sequence 62, Appl
75		18	1.1	54	3	US-09-479-645A-213	Sequence 213, App

ALIGNMENTS

RESULT 1
US-09-381-656-2
; Sequence 2, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11

```
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1590
;   TYPE: DNA
;   ORGANISM: Human
US-09-381-656-2

  Query Match             96.8%;  Score 1539;  DB 3;  Length 1590;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1589;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560

Qy       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590
              |||||||||||||||||||||||||||||
Db       1561 CCCCAAGGAGAGTTACTCGACAGTCCATAA 1590


RESULT 2
US-10-164-230-1
; Sequence 1, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 5337
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(99)
;   FEATURE:
```

```
;   NAME/KEY: exon
;   LOCATION: (2904)...(5337)
US-10-164-230-1

  Query Match             85.2%;  Score 1355;  DB 3;  Length 5337;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2901 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 2960

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2961 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 3020

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3021 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 3080

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3081 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 3140

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3141 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 3200

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3201 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 3260

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3261 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 3320

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       3321 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 3380

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3381 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 3440

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3441 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 3500

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3501 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 3560

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3561 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 3620

Qy        803 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3621 TTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 3680

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3681 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 3740

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 3800

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 3860

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 3920

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 3980

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 4040

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 4100

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 4160

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 4220

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 4280

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 4340

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4341 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 4400

Qy       1583 GTCCATAA 1590
              ||||||||
Db       4401 GTCCATAA 4408


RESULT 3
US-10-164-230-2
; Sequence 2, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 25235
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1)...(420)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1282)...(1405)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1702)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1602)...(1631)
```

```
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (2352)...(2364)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (6287)...(6509)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (10417)...(10493)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14407)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (14244)...(14344)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (25190)...(25235)
US-10-164-230-2

  Query Match             85.2%;  Score 1355;  DB 3;  Length 25235;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         83 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22552 CAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTA 22611

Qy        143 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22612 TCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCT 22671

Qy        203 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22672 ACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCT 22731

Qy        263 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22732 CCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGC 22791

Qy        323 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTACTCTACAAGGTG 382
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      22792 CAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCTAGTCTACAAGGTG 22851

Qy        383 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22852 CATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCACTCGGGAAGCAGCGGCTCTCACT 22911

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22912 CGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCAGCT 22971

Qy        503 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTAGCGCGGAATACTAA 562
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      22972 TCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGACAACTCTTACCGCGGAATACTAA 23031

Qy        563 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23032 ACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGGGTATCCAGCAGTG 23091

Qy        623 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23092 GCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCCGACTCTCCCTGCA 23151

Qy        683 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23152 GTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCACTCTGTGTCAG 23211

Qy        743 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23212 GGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCCCCTGGAGTGG 23271

Qy        803 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23272 TTCAAGGTCCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCT 23331

Qy        863 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23332 CTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGG 23391

Qy        923 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23392 TTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAAGAGA 23451

Qy        983 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23452 ACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGG 23511

Qy       1043 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23512 GCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCTTCCTCGGCCATTG 23571

Qy       1103 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23572 CGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCA 23631

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCG 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23632 AGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCT 23691

Qy       1223 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23692 CCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCA 23751

Qy       1283 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 1342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23752 CCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTT 23811

Qy       1343 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23812 GTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCCTCCTTGACACTGA 23871

Qy       1403 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23872 CTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCT 23931

Qy       1463 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23932 CCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGC 23991

Qy       1523 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23992 CTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACA 24051

Qy       1583 GTCCATAA 1590
              ||||||||
Db      24052 GTCCATAA 24059


RESULT 4
US-10-164-230-60
; Sequence 60, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 60
;   LENGTH: 1725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (15)...(1601)
US-10-164-230-60

  Query Match             60.8%;  Score 966;  DB 3;  Length 1725;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1566;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 74

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 134

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 194

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 254

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 314

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 374

Qy        361 TCTGGCTCTACTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 420
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TCTGGCTCTAGTCTACAAGGTGCATCCGGTTCCTCCCAGCTGGGGAGCAGCAGCTCTCAC 434

Qy        421 TCGGGAAGCAGCGGCTCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||| |||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db        435 TCGGGAARCAGCGGCTCTCACTCGGGAAGCYGCAGSTCTCATTCGAGCAGCAGCAGCAGC 494

Qy        481 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TTTCAGTTCAGCAGCAGCAGCTTCCAAGTAGGGAATGGCTCTGCTCTGCCAACCAATGAC 554

Qy        541 AACTCTTAGCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 600
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 AACTCTTACCGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAA 614

Qy        601 ACCTCTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 660
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 ACCTYTGGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCG 674

Qy        661 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GACATCCCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATC 734

Qy        721 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CCCAGCTCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCAC 794

Qy        781 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 GGTTCTGGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGC 854

Qy        841 AAGCCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 900
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 AAGWCCTGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGT 914

Qy        901 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 GGCTCCTCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCT 974

Qy        961 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTGGGCTACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCA 1034

Qy       1021 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 GCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAG 1094

Qy       1081 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TCGGCAGCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGT 1154

Qy       1141 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1214

Qy       1201 AGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1260
              |||||||||||||||||| ||| || ||||||||||||||||||||||||||||||||||
Db       1215 RGTTCTAGCATTTCCAGCAGCKCCGKTTYACCCTACCATCCCTGCGGCAGTGCTTCCCAG 1274

Qy       1261 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1334

Qy       1321 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATG 1380
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1335 GGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCARCTCTTCTGGTCACCCTTGCATG 1394

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1395 TCTGTCTCCTCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCC 1454

Qy       1441 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1455 GCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGG 1514

Qy       1501 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1515 GATATCCTAGCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTA 1574

Qy       1561 CCCCAAGGAGAGTTACTC 1578
              ||||||||||||||||||
Db       1575 CCCCAAGGAGAGTTACTC 1592


RESULT 5
US-10-164-230-15/c
; Sequence 15, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificially synthesized primer sequence
US-10-164-230-15
```

```
  Query Match             1.4%;  Score 22;  DB 3;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 4.6;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1537 CTAGCTGACCCTGAAGTTTTCC 1558
             ||||||||||||||||||||||
Db         22 CTAGCTGACCCTGAAGTTTTCC 1


RESULT 6
US-10-164-230-14
; Sequence 14, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificially synthesized primer sequence
US-10-164-230-14

  Query Match             1.3%;  Score 21;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1045 AAATACTTCTCCAGCAACCCC 1065
             |||||||||||||||||||||
Db          1 AAATACTTCTCCAGCAACCCC 21


RESULT 7
US-09-641-931F-35
; Sequence 35, Application US/09641931F
; Patent No. 6830885
; GENERAL INFORMATION:
;  APPLICANT: Lanctot, et al.
;  TITLE OF INVENTION: Nucleic Acid Molecule, Method and Kit for Selecting a
;  TITLE OF INVENTION:  Nucleic Acid Having A Desired Feature
;  FILE REFERENCE: 2003390-0001
;  CURRENT APPLICATION NUMBER: US/09/641,931F
;  CURRENT FILING DATE:  2003-04-08
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
;   LENGTH: 222
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-641-931F-35

  Query Match             1.3%;  Score 21;  DB 3;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        118 GACCCCTGTAAGGACCCCACG 138
             |||||||||||||||||||||
Db        132 GACCCCTGTAAGGACCCCACG 152


RESULT 8
US-09-248-796A-13896/c
; Sequence 13896, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13896
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-13896

  Query Match             1.3%;  Score 21;  DB 3;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GTGGTGGTGGTGGTGGACCAG 777
             |||||||||||||||||||||
Db        175 GTGGTGGTGGTGGTGGACCAG 155


RESULT 9
US-09-949-016-12030
; Sequence 12030, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12030
;   LENGTH: 56551
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12030

  Query Match             1.3%;  Score 21;  DB 3;  Length 56551;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 TCTGTAGACAAATCCTATGGT 882
             |||||||||||||||||||||
Db      41003 TCTGTAGACAAATCCTATGGT 41023


RESULT 10
US-09-949-016-12455/c
; Sequence 12455, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
```

```
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12455
;   LENGTH: 69813
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12455

  Query Match             1.3%;  Score 21;  DB 3;  Length 69813;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 GGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||
Db      58579 GGCCTGTGGTGGTGGTGGTGG 58559


RESULT 11
US-09-949-016-13905/c
; Sequence 13905, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13905
;   LENGTH: 69813
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13905

  Query Match             1.3%;  Score 21;  DB 3;  Length 69813;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 GGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||
Db      58579 GGCCTGTGGTGGTGGTGGTGG 58559


RESULT 12
US-09-949-016-13906/c
; Sequence 13906, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13906
;   LENGTH: 69813
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13906

  Query Match             1.3%;  Score 21;  DB 3;  Length 69813;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 GGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||
Db      58579 GGCCTGTGGTGGTGGTGGTGG 58559


RESULT 13
US-09-949-016-12861/c
; Sequence 12861, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12861
;   LENGTH: 69833
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12861

  Query Match             1.3%;  Score 21;  DB 3;  Length 69833;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 GGCCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||||
Db      58599 GGCCTGTGGTGGTGGTGGTGG 58579


RESULT 14
US-10-164-230-5/c
; Sequence 5, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 20
```

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificially synthesized primer sequence
US-10-164-230-5

  Query Match             1.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         61 CTGCTGGCTGGTCTCCTCCT 80
              ||||||||||||||||||||
Db         20 CTGCTGGCTGGTCTCCTCCT 1


RESULT 15
US-10-164-230-10
; Sequence 10, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificially synthesized primer sequence
US-10-164-230-10

  Query Match             1.3%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        297 GGGTGGTTCTGCAGGATCTT 316
              ||||||||||||||||||||
Db          1 GGGTGGTTCTGCAGGATCTT 20


RESULT 16
US-09-949-002-1172
; Sequence 1172, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1172
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-1172

  Query Match             1.3%;  Score 20;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCC 1179
              ||||||||||||||||||||
Db         72 CCAAGGGACCCTGCTCTCCC 91


RESULT 17
US-09-949-002-7268
; Sequence 7268, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7268
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-7268

  Query Match             1.3%;  Score 20;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCC 1179
              ||||||||||||||||||||
Db         72 CCAAGGGACCCTGCTCTCCC 91


RESULT 18
US-08-423-441-1/c
; Sequence 1, Application US/08423441
; Patent No. 5529926
;  GENERAL INFORMATION:
;    APPLICANT:  MAAT, JAN
;    APPLICANT:  MUSTERS, WOUTER
;    APPLICANT:  STAM, HEIN
;    APPLICANT:  SCHAAP, PETER J.
;    APPLICANT:  VAN DE VONDERVOORT, PETER J.
;    APPLICANT:  VISSER, JACOB
;    APPLICANT:  VERBAKEL, JOHANNES M.
;    TITLE OF INVENTION:  CLONING AND EXPRESSION OF DNA ENCODING A
;    TITLE OF INVENTION:  RIPENING FORM OF A POLYPEPTIDE HAVING SULFHYDRYL OXIDASE
;    TITLE OF INVENTION:  ACTIVITY
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/423,441
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/044,620
;      FILING DATE:  09-APR-1993
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, PAUL N.
;      REGISTRATION NUMBER:  16773
;      REFERENCE/DOCKET NUMBER:  202744/T7019(V)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3000
;      TELEFAX:  202-822-0944
;      TELEX:  248453CUSH
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2563 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Aspergillus niger
;      STRAIN:  CBS 120.49
;    IMMEDIATE SOURCE:
;      CLONE:  pUR7500, pUR7501
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  477..702
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  703..769
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  770..1573
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  1574..1681
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1682..1830
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(477..702, 770..1573, 1682..1830)
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  477..533
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1..476
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  1831..2563
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  join(534..702, 770..15731682..1827)
;      OTHER INFORMATION:  /product= "Sulfhydryl oxidase"
;      OTHER INFORMATION:  /gene= "SOX"
US-08-423-441-1

  Query Match             1.3%;  Score 20;  DB 2;  Length 2563;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        466 AGCAGCAGCAGCAGCTTTCA 485
              ||||||||||||||||||||
Db        391 AGCAGCAGCAGCAGCTTTCA 372


RESULT 19
US-09-949-002-580
; Sequence 580, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 580
;   LENGTH: 40325
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-580

  Query Match             1.3%;  Score 20;  DB 3;  Length 40325;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCC 1179
              ||||||||||||||||||||
Db      35059 CCAAGGGACCCTGCTCTCCC 35078


RESULT 20
US-09-949-002-771
; Sequence 771, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 771
;   LENGTH: 40325
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-771

  Query Match             1.3%;  Score 20;  DB 3;  Length 40325;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1160 CCAAGGGACCCTGCTCTCCC 1179
              ||||||||||||||||||||
Db      35059 CCAAGGGACCCTGCTCTCCC 35078


RESULT 21
US-10-164-230-12
; Sequence 12, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 19
```

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificially synthesized primer sequence
US-10-164-230-12

  Query Match             1.2%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        618 CAGTGGCCAAAGCGTCAGC 636
              ||||||||||||||||||||
Db          1 CAGTGGCCAAAGCGTCAGC 19


RESULT 22
US-09-023-942A-19
; Sequence 19, Application US/09023942A
; Patent No. 6479274
;  GENERAL INFORMATION:
;    APPLICANT:  (US only) ANTALIS Toni Marie and HOOPER John David
;    TITLE OF INVENTION:  NOVEL MOLECULES
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,942A
;      FILING DATE:  13-FEB-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PO5101/97
;      FILING DATE:  13-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP0422/97
;      FILING DATE:  18-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  International PCT Application
;      FILING DATE:  13-FEB-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DIGIGLIO,  FRANK S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  11168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742 4343
;      TELEFAX:  (516) 742 4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-023-942A-19

  Query Match             1.2%;  Score 19;  DB 3;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GTGGTGGTGGTGGTGGACC 775
              ||||||||||||||||||||
Db         18 GTGGTGGTGGTGGTGGACC 36


RESULT 23
US-08-348-143-14
; Sequence 14, Application US/08348143
; Patent No. 5506205
;  GENERAL INFORMATION:
;    APPLICANT:  TAJIMA, HISAO
;    APPLICANT:  KITAGAWA, KOICHIRO
;    APPLICANT:  OHNO, HIROYUKI
;    APPLICANT:  UENO, TOSHIO
;    TITLE OF INVENTION:  A No. 5506205el Polypeptide of Protein p140 and DNAs
;    TITLE OF INVENTION:  encoding it
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/348,143
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 315806/1993
;      FILING DATE:  24-NOV-1993
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (synthetic)
US-08-348-143-14

  Query Match             1.2%;  Score 19;  DB 2;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GTGGTGGTGGTGGTGGACC 775
              ||||||||||||||||||||
Db         19 GTGGTGGTGGTGGTGGACC 37


RESULT 24
US-08-571-785-14
; Sequence 14, Application US/08571785
; Patent No. 5804411
;  GENERAL INFORMATION:
;    APPLICANT:  TAJIMA, HISAO
;    APPLICANT:  KITAGAWA, KOICHIRO
;    APPLICANT:  OHNO, HIROYUKI
;    APPLICANT:  UENO, TOSHIO
;    TITLE OF INVENTION:  A No. 5804411el Polypeptide of Protein p140
;    TITLE OF INVENTION:  and DNAs encoding it
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037-3202
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/571,785
;      FILING DATE:  13-DEC-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/348,143
;      FILING DATE:  23-NOV-1994
;      APPLICATION NUMBER:  JP 315806/1993
;      FILING DATE:  24-NOV-1993
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (synthetic)
US-08-571-785-14

  Query Match             1.2%;  Score 19;  DB 2;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GTGGTGGTGGTGGTGGACC 775
              |||||||||||||||||||
Db         19 GTGGTGGTGGTGGTGGACC 37


RESULT 25
US-09-192-435-14
; Sequence 14, Application US/09192435
; Patent No. 6303320
;  GENERAL INFORMATION:
;    APPLICANT:  TAJIMA, HISAO
;    APPLICANT:  KITAGAWA, KOICHIRO
;    APPLICANT:  OHNO, HIROYUKI
;    APPLICANT:  UENO, TOSHIO
;    TITLE OF INVENTION:  A No. 6303320el Polypeptide of Protein p140
;    TITLE OF INVENTION:  and DNAs encoding it
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/192,435
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/571,785
;      FILING DATE:  13-DEC-1995
;      APPLICATION NUMBER:  08/348,143
;      FILING DATE:  23-NOV-1994
;      APPLICATION NUMBER:  JP 315806/1993
;      FILING DATE:  24-NOV-1993
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (synthetic)
US-09-192-435-14

  Query Match             1.2%;  Score 19;  DB 3;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        757 GTGGTGGTGGTGGTGGACC 775
              |||||||||||||||||||
Db         19 GTGGTGGTGGTGGTGGACC 37


RESULT 26
US-09-558-340-14
; Sequence 14, Application US/09558340
; Patent No. 6432913
;  GENERAL INFORMATION:
;    APPLICANT:  TAJIMA, HISAO
;    APPLICANT:  KITAGAWA, KOICHIRO
;    APPLICANT:  OHNO, HIROYUKI
;    APPLICANT:  UENO, TOSHIO
;    TITLE OF INVENTION:  A No. 6432913el Polypeptide of Protein p140
;    TITLE OF INVENTION:  and DNAs encoding it
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/558,340
;      FILING DATE:  26-APR-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/192,435
;      FILING DATE:  08-JAN-1998
;      APPLICATION NUMBER:  08/571,785
;      FILING DATE:  13-DEC-1995
;      APPLICATION NUMBER:  08/348,143
;      FILING DATE:  23-NOV-1994
;      APPLICATION NUMBER:  JP 315806/1993
;      FILING DATE:  24-NOV-1993
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (synthetic)
US-09-558-340-14

  Query Match             1.2%;  Score 19;  DB 3;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        757 GTGGTGGTGGTGGTGGACC 775
              ||||||||||||||||||
Db         19 GTGGTGGTGGTGGTGGACC 37


RESULT 27
US-09-270-767-8813
; Sequence 8813, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8813
;   LENGTH: 233
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-8813

  Query Match             1.2%;  Score 19;  DB 3;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        462 TTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||
Db        191 TTCGAGCAGCAGCAGCAGC 209


RESULT 28
US-09-270-767-24095
; Sequence 24095, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24095
;   LENGTH: 233
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-24095

  Query Match             1.2%;  Score 19;  DB 3;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        462 TTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||
Db        191 TTCGAGCAGCAGCAGCAGC 209


RESULT 29
US-08-209-747-11/c
; Sequence 11, Application US/08209747
; Patent No. 5733771
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Colgin, Mark
;    TITLE OF INVENTION:  cDNAs Encoding Minor Ampullate Spider
;    TITLE OF INVENTION:  Silk Proteins
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,747
;      FILING DATE:  14-MAR-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-104P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  N. clavipes
;      TISSUE TYPE:  minor ampullate gland
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..240
;      OTHER INFORMATION:  /label= cloned_cDNA
;      OTHER INFORMATION:  /note= "partial sequence of pMISS3, 11-1 template,
;      OTHER INFORMATION:  reverse primer"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..240
;      OTHER INFORMATION:  /product= "pMISS3 partial sequence
;      OTHER INFORMATION:  translation"
US-08-209-747-11

  Query Match             1.2%;  Score 19;  DB 2;  Length 240;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1293 CAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||
Db        166 CAGCAGCAGCTCCAGTTCC 148


RESULT 30
US-08-458-298-11/c
; Sequence 11, Application US/08458298
; Patent No. 5756677
;  GENERAL INFORMATION:
;    APPLICANT:  Lewis, Randolph V.
;    APPLICANT:  Colgin, Mark
;    TITLE OF INVENTION:  cDNAs Encoding Minor Ampullate Spider
;    TITLE OF INVENTION:  Silk Proteins
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,298
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,747
;      FILING DATE:  14-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1447-104P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  N. clavipes
;      TISSUE TYPE:  minor ampullate gland
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..240
;      OTHER INFORMATION:  /label= cloned_cDNA
;      OTHER INFORMATION:  /note= "partial sequence of pMISS3, 11-1 template,
;      OTHER INFORMATION:  reverse primer"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..240
;      OTHER INFORMATION:  /product= "pMISS3 partial sequence
;      OTHER INFORMATION:  translation"
US-08-458-298-11

  Query Match             1.2%;  Score 19;  DB 2;  Length 240;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1293 CAGCAGCAGCTCCAGTTCC 1311
             |||||||||||||||||||
Db        166 CAGCAGCAGCTCCAGTTCC 148


RESULT 31
US-09-854-133-659
; Sequence 659, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 659
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-854-133-659

  Query Match             1.2%;  Score 19;  DB 3;  Length 366;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        483 TCAGTTCAGCAGCAGCAGC 501
              |||||||||||||||||||
Db        192 TCAGTTCAGCAGCAGCAGC 210


RESULT 32
US-09-621-976-14545
; Sequence 14545, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 14545
;   LENGTH: 523
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 501
;   OTHER INFORMATION: n=a, g, c or t
US-09-621-976-14545

  Query Match             1.2%;  Score 19;  DB 3;  Length 523;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        748 CAGAGGCCTGTGGTGGTGG 766
              |||||||||||||||||||
Db        203 CAGAGGCCTGTGGTGGTGG 221


RESULT 33
US-09-949-016-61662/c
; Sequence 61662, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61662
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-61662

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        754 CCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||
Db        402 CCTGTGGTGGTGGTGGTGG 384
```

```
RESULT 34
US-09-949-016-61784/c
; Sequence 61784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61784
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-61784

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        754 CCTGTGGTGGTGGTGGTGG 772
              |||||||||||||||||||
Db        402 CCTGTGGTGGTGGTGGTGG 384


RESULT 35
US-09-949-016-94504
; Sequence 94504, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94504
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-94504

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        972 CACCAAAGAGAACCCTGTG 990
              |||||||||||||||||||
Db        242 CACCAAAGAGAACCCTGTG 260


RESULT 36
US-09-949-016-94505
; Sequence 94505, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94505
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-94505

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        972 CACCAAAGAGAACCCTGTG 990
              |||||||||||||||||||
Db        417 CACCAAAGAGAACCCTGTG 435


RESULT 37
US-09-949-016-94506
; Sequence 94506, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94506
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-94506

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        972 CACCAAAGAGAACCCTGTG 990
              |||||||||||||||||||
Db        502 CACCAAAGAGAACCCTGTG 520


RESULT 38
US-09-949-002-1670
; Sequence 1670, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
```

```
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1670
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-1670

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1442 CTGGTGCCAAGCCCTGTGG 1460
             ||||||||||||||||||||
Db        386 CTGGTGCCAAGCCCTGTGG 404


RESULT 39
US-09-949-002-1671
; Sequence 1671, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1671
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-1671

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1442 CTGGTGCCAAGCCCTGTGG 1460
             ||||||||||||||||||||
Db        502 CTGGTGCCAAGCCCTGTGG 520


RESULT 40
US-09-949-002-5678
; Sequence 5678, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5678
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-5678

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1442 CTGGTGCCAAGCCCTGTGG 1460
             ||||||||||||||||||||
Db        386 CTGGTGCCAAGCCCTGTGG 404


RESULT 41
US-09-949-002-5679
; Sequence 5679, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5679
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-5679

  Query Match             1.2%;  Score 19;  DB 3;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1442 CTGGTGCCAAGCCCTGTGG 1460
             ||||||||||||||||||||
Db        502 CTGGTGCCAAGCCCTGTGG 520


RESULT 42
US-09-248-796A-499/c
; Sequence 499, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 499
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-499

  Query Match             1.2%;  Score 19;  DB 3;  Length 750;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        489 CAGCAGCAGCAGCTTCCAA 507
              ||||||||||||||||||||
Db        658 CAGCAGCAGCAGCTTCCAA 640


RESULT 43
```

```
US-09-270-767-5649
; Sequence 5649, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5649
;   LENGTH: 751
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: n means any nucleotide
US-09-270-767-5649

  Query Match             1.2%;  Score 19;  DB 3;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        462 TTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||
Db        724 TTCGAGCAGCAGCAGCAGC 742


RESULT 44
US-09-270-767-20931
; Sequence 20931, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20931
;   LENGTH: 751
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: n means any nucleotide
US-09-270-767-20931

  Query Match             1.2%;  Score 19;  DB 3;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        462 TTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||
Db        724 TTCGAGCAGCAGCAGCAGC 742


RESULT 45
US-08-998-416-218
; Sequence 218, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  218:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  777 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1205UP
US-08-998-416-218

  Query Match             1.2%;  Score 19;  DB 3;  Length 777;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         67 GCTGGTCTCCTCCTGCCAG 85
              ||||||||||||||||||||
Db        244 GCTGGTCTCCTCCTGCCAG 262


RESULT 46
PCT-US95-03323A-1/c
; Sequence 1, Application PC/TUS9503323A
;  GENERAL INFORMATION:
;    APPLICANT:  KUNSCH, ET AL.
;    TITLE OF INVENTION:  Human B-Cell Translocation Genes-2 and 3
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03323A
;      FILING DATE:  Submitted herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-262
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1843 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
PCT-US95-03323A-1

  Query Match             1.2%;  Score 19;  DB 6;  Length 1843;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        754 CCTGTGGTGGTGGTGGTGG 772
              ||||||||||||||||||||
Db        798 CCTGTGGTGGTGGTGGTGG 780


RESULT 47
US-09-799-451-477
; Sequence 477, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Reena
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6783969el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 803
;  CURRENT APPLICATION NUMBER: US/09/799,451
;  CURRENT FILING DATE:  2001-03-05
;  NUMBER OF SEQ ID NOS: 948
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 477
;   LENGTH: 2565
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (169)..(1140)
US-09-799-451-477

  Query Match             1.2%;  Score 19;  DB 3;  Length 2565;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        462 TTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||
Db        306 TTCGAGCAGCAGCAGCAGC 324


RESULT 48
US-09-830-230A-629/c
; Sequence 629, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Lyme Disease Vaccines
;  FILE REFERENCE: PB481US
;  CURRENT APPLICATION NUMBER: US/09/830,230A
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: PCT/US98/12718
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: 60/057,483
;  PRIOR FILING DATE: 1997-09-03
;  PRIOR APPLICATION NUMBER: 60/053,344
;  PRIOR FILING DATE: 1997-07-22
;  PRIOR APPLICATION NUMBER: 60/053,377
;  PRIOR FILING DATE: 1997-07-22
;  PRIOR APPLICATION NUMBER: 60/050,359
;  PRIOR FILING DATE: 1997-06-20
;  NUMBER OF SEQ ID NOS: 756
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 629
;   LENGTH: 3990
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1139)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1143)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1148)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1210)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1244)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1247)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1250)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1251)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-830-230A-629

  Query Match             1.2%;  Score 19;  DB 3;  Length 3990;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        468 CAGCAGCAGCAGCTTTCAG 486
              ||||||||||||||||||||
Db       3874 CAGCAGCAGCAGCTTTCAG 3856


RESULT 49
US-08-146-930-1
; Sequence 1, Application US/08146930
; Patent No. 5958764
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Roop, Dennis R.
;    APPLICANT:  Rothnagel, Joseph A.
;    APPLICANT:  Greenhalgh, David A.
;    TITLE OF INVENTION:  SPECIFIC EXPRESSION VECTORS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LYON & LYON
;      STREET:  611 West Sixth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90017
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/146,930
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:             two
;      APPLICATION NUMBER:  07/876,286
;      FILING DATE:  April 30, 1992
;      APPLICATION NUMBER:  No. 5958764 yet assigned (204/132)
;      FILING DATE:  October 29, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  204/152
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6530 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-146-930-1

  Query Match             1.2%;  Score 19;  DB 2;  Length 6530;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        249 CTCTGGTGGTGGCTCCAGT 267
              ||||||||||||||||||||
Db       3638 CTCTGGTGGTGGCTCCAGT 3656


RESULT 50
US-08-458-240-1
; Sequence 1, Application US/08458240
; Patent No. 6143727
;  GENERAL INFORMATION:
;    APPLICANT:  Roop, Dennis R.
;    APPLICANT:  Rothnagel, Joseph A.
;    APPLICANT:  Greenhalgh, David A.
;    TITLE OF INVENTION:  SPECIFIC EXPRESSION VECTORS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LYON & LYON
;      STREET:  611 West Sixth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90017
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  IBM MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,240
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/146,930
;      FILING DATE:
;      APPLICATION NUMBER:  07/876,286
;      FILING DATE:  April 30, 1992
;      APPLICATION NUMBER:  No. 6143727 yet assigned (204/132)
;      FILING DATE:  October 29, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  204/152
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6530 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-458-240-1

  Query Match             1.2%;  Score 19;  DB 3;  Length 6530;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        249 CTCTGGTGGTGGCTCCAGT 267
              ||||||||||||||||||||
Db       3638 CTCTGGTGGTGGCTCCAGT 3656


Search completed: March 22, 2006, 03:06:46
Job time : 285 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:20:04 ; Search time 6.8 Seconds
(without alignments)
4499.549 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565918

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152106_26838/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=50
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_25_@runat_21032006_152106_26838 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	224	42.4	502	2	A48679	differentiated ker
2	12	2.3	733	2	S44876	ZC21.4 protein - C
3	11	2.1	698	2	A84547	hypothetical prote
4	10	1.9	206	2	T23852	hypothetical prote
5	10	1.9	344	2	T40167	hypothetical prote
6	10	1.9	365	2	A49235	outer membrane lip
7	10	1.9	375	2	D87427	conserved hypothet
8	10	1.9	710	2	S28014	outD protein - Erw
9	9	1.7	167	2	T47385	hypothetical prote
10	9	1.7	277	2	D84596	hypothetical prote
11	9	1.7	299	2	T12483	hypothetical prote
12	9	1.7	345	2	B43731	achaete-scute comp
13	9	1.7	356	2	JX0141	cytochrome-c oxida
14	9	1.7	357	2	F96708	hypothetical prote
15	9	1.7	367	2	I39649	lipoprotein - Acti
16	9	1.7	406	2	G96497	hypothetical prote
17	9	1.7	427	2	C88883	protein JC8.10 [im
18	9	1.7	427	2	A32372	female-specific do
19	9	1.7	441	2	G01758	transcription fact
20	9	1.7	450	2	C95348	HemN coproporphyri
21	9	1.7	452	2	T40634	hypothetical prote
22	9	1.7	453	2	T48240	hypothetical prote
23	9	1.7	465	2	F75524	hypothetical prote
24	9	1.7	474	2	I38240	transcription fact
25	9	1.7	510	2	T37541	probable glycolipi
26	9	1.7	549	2	B32372	male-specific doub
27	9	1.7	623	1	S33167	gene pointed prote
28	9	1.7	644	2	S39356	transcription fact
29	9	1.7	683	2	A82704	1,4-beta-cellobios
30	9	1.7	837	2	D84428	hypothetical prote
31	9	1.7	1065	2	T13230	dachshund isoform
32	9	1.7	1072	2	T13232	dachshund protein
33	9	1.7	1074	2	T13229	dachshund protein
34	9	1.7	1081	2	T13231	dachshund protein
35	9	1.7	1139	2	B54962	sterol regulatory
36	9	1.7	1337	2	T38949	hypothetical prote
37	9	1.7	1428	2	T08852	lustrin A - Califo
38	9	1.7	1469	2	T09219	basal transcriptio
39	9	1.7	1785	2	A45546	major merozoite su
40	9	1.7	3972	2	S75251	hypothetical prote
c 41	8	1.5	89	2	T45015	hypothetical prote
c 42	8	1.5	97	1	ERBP1Z	gene 1C protein -
43	8	1.5	99	2	D86931	hypothetical prote
44	8	1.5	100	2	A72572	hypothetical prote
45	8	1.5	115	2	D61615	sericin MG-1 - gre
46	8	1.5	120	2	T45451	hypothetical prote
47	8	1.5	120	2	A45519	variant surface gl
48	8	1.5	125	2	S69871	hypothetical prote
49	8	1.5	129	2	T00282	hypothetical prote
50	8	1.5	141	2	A86417	probable auxin-ind

ALIGNMENTS

RESULT 1
A48679
differentiated keratinocyte S protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A48679
R;Zhou, Y.; Chaplin, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 9470-9474, 1993
A;Title: Identification in the HLA class I region of a gene expressed late in keratinoc
A;Reference number: A48679; MUID:94022396; PMID:8415725
A;Accession: A48679
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <ZHO>
A;Cross-references: UNIPROT:O43509; UNIPARC:UPI0000177376; GB:L20815
A;Note: sequence shown follows the authors' translation at positions 47 and 410
C;Genetics:
A;Map position: 6p21.3
C;Superfamily: loricrin

Alignment Scores:

Pred. No.:	2.02e-205	Length:	502
Score:	224.00	Matches:	224
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	42.4%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x A48679 (1-502)

Qy 550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
Db 184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

```
Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        404 IleSerSerSer 407
```

RESULT 2
S44876
ZC21.4 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C;Accession: S44876
R;Du, Z.; Waterston, R.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid ZC21.
A;Reference number: S44649
A;Accession: S44876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-733 <DUZ>
A;Cross-references: UNIPARC:UPI000017BD58; EMBL:L16685; NID:g289729; PID:g289735
C;Genetics:
A;Introns: 269/3; 551/3; 600/2; 670/3

Alignment Scores:

Pred. No.:	0.0129	Length:	733
Score:	12.00	Matches:	12
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.3%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x S44876 (1-733)

```
Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              |||||||||||||||||||||||||||||||||||||
Db         50 SerGlyGlyGlySerSerGlySerSerSerGlySer 61
```

RESULT 3
A84547
hypothetical protein At2g17010 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84547
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84547
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-698 <STO>
A;Cross-references: UNIPROT:Q7XJL9; UNIPARC:UPI000016233B; GB:AE002093; NID:g6598358; P
C;Genetics:
A;Gene: At2g17010
A;Map position: 2

Alignment Scores:

Pred. No.:	0.118	Length:	698
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x A84547 (1-698)

```
Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||||||||
Db        196 SerSerSerSerHisSerSerSerSerSerSer 206
```

RESULT 4
T23852
hypothetical protein R02D5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23852
R;Matthews, L.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19808
A;Accession: T23852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-206 <WIL>
A;Cross-references: UNIPROT:Q21636; UNIPARC:UPI000007FF99; EMBL:Z78015; PIDN:CAB01433.1
A;Experimental source: clone R02D5
C;Genetics:
A;Gene: CESP:R02D5.1
A;Map position: 5
A;Introns: 53/3

Alignment Scores:

Pred. No.:	1.32	Length:	206
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T23852 (1-206)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db         84 SerSerSerHisSerSerSerSerSerSer 93
```

RESULT 5
T40167
hypothetical protein SPBC30B4.01c - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40167; T40374
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, August 1998
A;Reference number: Z21909
A;Accession: T40167
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-344 <LYN>
A;Cross-references: UNIPROT:P87179; UNIPARC:UPI000013A1DF; EMBL:AL031262; PIDN:CAA20314.
A;Experimental source: strain 972h-; cosmid c30B4
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21923
A;Accession: T40374
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 36-344 <WOO>
A;Cross-references: UNIPARC:UPI00001620C6; EMBL:Z95620; PIDN:CAB09116.1; GSPDB:GN00067;
A;Experimental source: strain 972h-; cosmid c3D6
C;Genetics:
A;Gene: SPBC3D5.14C; SPDB:SPBC30B4.01c; SPDB:SPBC3D6.14c
A;Map position: 2

Alignment Scores:

Pred. No.:	1.21	Length:	344
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T40167 (1-344)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        185 SerSerSerHisSerSerSerSerSerSer 194
```

RESULT 6
A49235
outer membrane lipoprotein OmlA - Actinobacillus pleuropneumoniae
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A49235
R;Gerlach, G.F.; Anderson, C.; Klashinsky, S.; Rossi-Campos, A.; Potter, A.A.; Willson,
Infect. Immun. 61, 565-572, 1993
A;Title: Molecular characterization of a protective outer membrane lipoprotein (OmlA) fr
A;Reference number: A49235; MUID:93138779; PMID:8423086
A;Accession: A49235
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-365 <GER>
A;Cross-references: UNIPROT:Q02937; UNIPARC:UPI0000130CE2
A;Experimental source: serotype 1
A;Note: sequence extracted from NCBI backbone (NCBIN:123233, NCBIP:123234)

Alignment Scores:

Pred. No.:	1.2	Length:	365
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x A49235 (1-365)

```
Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30
```

RESULT 7
D87427
conserved hypothetical protein CC1435 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: D87427
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87427
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-375 <STO>
A;Cross-references: UNIPROT:Q9A8B9; UNIPARC:UPI00000C73AF; GB:AE005673; NID:g13422800;
C;Genetics:
A;Gene: CC1435

Alignment Scores:

Pred. No.:	1.19	Length:	375
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x D87427 (1-375)

```
Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCT 481
              ||||||||||||||||||||||||||||||
Db        121 AlaAlaLeuIleArgAlaAlaAlaAlaAla 130
```

RESULT 8
S28014
outD protein - Erwinia chrysanthemi
C;Species: Erwinia chrysanthemi
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: S28014; S23886
R;Condemine, G.; Dorel, C.; Hugouvieux-Cotte-Pattat, N.; Robert-Baudouy, J.
Mol. Microbiol. 6, 3199-3211, 1992
A;Title: Some of the out genes involved in the secretion of pectate lyases in Erwinia c
A;Reference number: S28011; MUID:93086427; PMID:1453958
A;Accession: S28014
A;Molecule type: DNA
A;Residues: 1-710 <CON>
A;Cross-references: UNIPROT:Q01565; UNIPARC:UPI000012BC7E; EMBL:X65265; NID:g3152953; P
C;Genetics:
A;Gene: outD
A;Start codon: GTG

Alignment Scores:

Pred. No.:	1.07	Length:	710
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x S28014 (1-710)

```
Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        321 GlySerSerGlySerSerSerGlySerSer 330
```

RESULT 9
T47385
hypothetical protein T5C2.130 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47385

R;Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24463
A;Accession: T47385
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-167 <OBE>
A;Cross-references: UNIPROT:Q9M176; UNIPARC:UPI00000A6935; EMBL:AL138664
A;Experimental source: cultivar Columbia; BAC clone T5C2
C;Genetics:
A;Map position: 3
A;Note: T5C2.130
F;79-139/Domain: RING finger homology <RRN>

Alignment Scores:

Pred. No.:	12.4	Length:	167
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T47385 (1-167)

```
Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||
Db         41 SerSerSerSerSerSerGlnSerSer 49
```

RESULT 10
D84596
hypothetical protein At2g21040 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84596
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84596
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-277 <STO>
A;Cross-references: UNIPROT:Q9SKQ7; UNIPARC:UPI000017A77F; GB:AE002093; NID:g4803939; PI
C;Genetics:
A;Gene: At2g21040
A;Map position: 2

Alignment Scores:

Pred. No.:	11.4	Length:	277
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x D84596 (1-277)

```
Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGC 273
              ||||||||||||||||||||||||||
Db        111 SerSerGlyGlyGlySerSerGlySer 119
```

RESULT 11
T12483
hypothetical protein DKFZp564B0769.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: T12483
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17527
A;Accession: T12483
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <ANS>
A;Cross-references: UNIPROT:Q9Y4N3; UNIPARC:UPI000006CCAE; EMBL:AL080186
A;Experimental source: fetal brain; clone DKFZp564B0769
C;Genetics:
A;Note: DKFZp564B0769.1

Alignment Scores:

Pred. No.:	11.2	Length:	299
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T12483 (1-299)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 12
B43731
achaete-scute complex protein T4 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C;Accession: B43731; S35425
R;Villares, R.; Cabrera, C.V.
Cell 50, 415-424, 1987
A;Title: The achaete-scute gene complex of Drosophila melanogaster: conserved domains i
A;Reference number: A43731; MUID:87273503; PMID:3111716
A;Accession: B43731
A;Molecule type: DNA
A;Residues: 1-345 <VIL1>
A;Cross-references: UNIPROT:P10084; UNIPARC:UPI000017BE72; GB:M17119
R;Villares, R.
submitted to the EMBL Data Library, November 1990
A;Reference number: S35425
A;Accession: S35425
A;Molecule type: DNA
A;Residues: 1-255,'C',257-345 <VIL2>
A;Cross-references: UNIPARC:UPI000016BA9C; EMBL:M17119; NID:g156745; PID:g156748
C;Genetics:
A;Gene: FlyBase:sc
A;Cross-references: FlyBase:FBgn0004170

Alignment Scores:

Pred. No.:	10.9	Length:	345
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x B43731 (1-345)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||
Db        191 SerSerSerHisSerSerSerSerSer 199
```

RESULT 13
JX0141
cytochrome-c oxidase (EC 1.9.3.1) chain II precursor - thermophilic bacterium PS-3
N;Alternate names: cytochrome a3; cytochrome aa3
C;Species: thermophilic bacterium PS-3
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C;Accession: JX0141; A38831
R;Ishizuka, M.; Machida, K.; Shimada, S.; Mogi, A.; Tsuchiya, T.; Ohmori, T.; Souma, Y.
J. Biochem. 108, 866-873, 1990
A;Title: Nucleotide sequence of the gene coding for four subunits of cytochrome c oxida
A;Reference number: JX0140; MUID:91185348; PMID:1964459
A;Accession: JX0141

A;Molecule type: DNA
A;Residues: 1-356 <ISH>
A;Cross-references: UNIPARC:UPI0000127FDB; GB:D13955; GB:D00728; NID:g217148; PIDN:BAA03
A;Accession: A38831
A;Molecule type: protein
A;Residues: 'X',25-48 <IS2>
A;Cross-references: UNIPARC:UPI0000175187
C;Genetics:
A;Gene: caaA
C;Superfamily: cytochrome-c oxidase, subunit II; cytochrome-c oxidase chain II homology
C;Keywords: copper; electron transfer; heme; membrane-associated complex; oxidoreductase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-356/Product: cytochrome-c oxidase chain II #status experimental <MAT>
F;219,221,223,227/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted
F;221/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Alignment Scores:

Pred. No.:	10.9	Length:	356
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x JX0141 (1-356)

```
Qy         46 GGGATGATGGCACTGCTGCTGGCTGGT 72
              |||||||||||||||||||||||||||
Db         15 GlyMetMetAlaLeuLeuLeuAlaGly 23
```

RESULT 14
F96708
hypothetical protein T26J14.3 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F96708
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <STO>
A;Cross-references: UNIPROT:Q9CA35; UNIPARC:UPI0000048424; GB:AE005173; NID:g6553935; PI
C;Genetics:
A;Gene: T26J14.3
A;Map position: 1
C;Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:

Pred. No.:	10.9	Length:	357
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x F96708 (1-357)

```
Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35
```

RESULT 15
I39649
lipoprotein - Actinobacillus pleuropneumoniae
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C;Accession: I39649; I39518; S52740
R;Bunka, S.; Christensen, C.; Potter, A.A.; Willson, P.J.; Gerlach, G.F.
Infect. Immun. 63, 2797-2800, 1995
A;Title: Cloning and characterization of a protective outer membrane lipoprotein of Act
A;Reference number: I39649; MUID:95310047; PMID:7790104
A;Accession: I39649
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-367 <RES>
A;Cross-references: UNIPROT:Q44164; UNIPARC:UPI00000B4E8C; EMBL:Z48920; NID:g757528; PI
R;Ito, H.; Uchida, I.; Sekizaki, T.; Ooishi, E.; Kawai, T.; Okabe, T.; Taneno, A.; Tera
Microb. Pathog. 18, 29-36, 1995
A;Title: Molecular cloning of an Actinobacillus pleuropneumoniae outer membrane lipoprc
A;Reference number: I39518; MUID:95302932; PMID:7783596
A;Accession: I39518
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-361,'KEDK',366 <RE2>
A;Cross-references: UNIPARC:UPI00000BB1F1; GB:D28491; NID:g808858; PIDN:BAA05852.1; PID
C;Genetics:
A;Gene: omlA

Alignment Scores:

Pred. No.:	10.8	Length:	367
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x I39649 (1-367)

```
Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCC 276
              ||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSer 29
```

RESULT 16
G96497
hypothetical protein F2H10.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C;Accession: G96497
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonsc
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96497
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <STO>
A;Cross-references: UNIPROT:Q9C8C6; UNIPARC:UPI00000AA4E0; GB:AE005173; NID:g11094798;
C;Genetics:
A;Gene: F2H10.4
A;Map position: 1
C;Superfamily: Putative ternary complex factor MIP1

Alignment Scores:

Pred. No.:	10.6	Length:	406
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

```
US-10-677-865A-2 (1-1590) x G96497 (1-406)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCCCAA 1314
              ||||||||||||||||||||||||||||
Db         13 SerPheSerSerSerSerSerSerGln 21

RESULT 17
C88883
protein JC8.10 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C88883
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: C88883
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <STO>
A;Cross-references: UNIPROT:Q9XUD3; UNIPROT:Q9GT42; UNIPARC:UPI000017A631; GB:chr_IV; PI
C;Genetics:
A;Gene: JC8.10
A;Map position: 4

Alignment Scores:
Pred. No.:              10.5          Length:          427
Score:                  9.00          Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.7%          Indels:          0
DB:                     2             Gaps:            0

US-10-677-865A-2 (1-1590) x C88883 (1-427)

Qy       1381 TCTGTCTCCTCCTTGACACTGACTGGG 1407
              ||||||||||||||||||||||||||
Db        291 SerValSerSerLeuThrLeuThrGly 299

RESULT 18
A32372
female-specific doublesex protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 09-Jul-2004
C;Accession: A32372
R;Burtis, K.C.; Baker, B.S.
Cell 56, 997-1010, 1989
A;Title: Drosophila doublesex gene controls somatic sexual differentiation by producing
A;Reference number: A32372; MUID:89168451; PMID:2493994
A;Accession: A32372
A;Status: preliminary
A;Molecule type: DNA; mRNA
A;Residues: 1-427 <BUR>
A;Cross-references: UNIPROT:P23023; UNIPARC:UPI000000751A; GB:M25292; GB:J03156
C;Genetics:
A;Gene: FlyBase:dsx
A;Cross-references: FlyBase:FBgn0000504
C;Keywords: alternative splicing; DNA binding; nucleus; transcription regulation

Alignment Scores:
Pred. No.:              10.5          Length:          427
Score:                  9.00          Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.7%          Indels:          0
DB:                     2             Gaps:            0

US-10-677-865A-2 (1-1590) x A32372 (1-427)

Qy        211 TCCAGCAGTTCTGGCAGCTCCATTTCC 237
              ||||||||||||||||||||||||||
Db         29 SerSerSerSerGlySerSerIleSer 37

RESULT 19
G01758
transcription factor SOX11 - human
N;Alternate names: Aip-1
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Oct-2004
C;Accession: G01758; JC7156
R;Jay, P.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08331
A;Accession: G01758
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-441 <JAY>
A;Cross-references: UNIPROT:P35716; UNIPARC:UPI00001362AD; EMBL:U23752; NID:g974247; PI
R;Azuma, T.; Ao, S.; Saito, Y.; Yano, K.; Seki, N.; Wakao, H.; Masuho, Y.; Muramatsu, M
DNA Res. 6, 357-360, 1999
A;Title: Human SOX11, an upregulated gene during the neural differentiation, has a long
A;Reference number: JC7156; MUID:20039622; PMID:10574465
A;Accession: JC7156
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <AZU>
A;Cross-references: UNIPARC:UPI00001362AD; DDBJ:AB028641; NID:g6527111; PIDN:BAA88122.1
C;Genetics:
A;Gene: GDB:SOX11; SOX-11
A;Cross-references: GDB:677487; OMIM:600898
A;Map position: 2p25-2p25
F;46-121/Domain: HMG box homology <HMG1>

Alignment Scores:
Pred. No.:              10.5          Length:          441
Score:                  9.00          Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.7%          Indels:          0
DB:                     2             Gaps:            0

US-10-677-865A-2 (1-1590) x G01758 (1-441)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||
Db        346 SerSerGlySerSerSerGlySerSer 354

RESULT 20
C95348
HemN coproporphyrinogen III oxidase [imported] - Sinorhizobium meliloti (strain 1021) m
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95348
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: C95348
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-450 <KUR>
A;Cross-references: UNIPROT:Q92Z19; UNIPARC:UPI00000CB189; GB:AE006469; PIDN:AAK65349.1
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
```

C;Genetics:
A;Gene: hemN
A;Genome: plasmid
C;Superfamily: oxygen-independent coproporphyrinogen oxidase

Alignment Scores:
Pred. No.: 10.4 Length: 450
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x C95348 (1-450)

Qy 458 CTCATTCGAGCAGCAGCAGCAGCTTTC 484
||||||||||||||||||||||||||||
Db 427 LeuIleArgAlaAlaAlaAlaAlaPhe 435

RESULT 21
T40634
hypothetical protein SPBC685.03 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40634
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21928
A;Accession: T40634
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-452 <SEE>
A;Cross-references: UNIPROT:Q9Y7L5; UNIPARC:UPI000006A6FA; EMBL:AL049474; PIDN:CAB39360.
A;Experimental source: strain 972h-; cosmid c685
C;Genetics:
A;Gene: SPDB:SPBC685.03
A;Map position: 2

Alignment Scores:
Pred. No.: 10.4 Length: 452
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T40634 (1-452)

Qy 1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
||||||||||||||||||||||||||||
Db 177 SerSerSerSerSerSerGlnSerSer 185

RESULT 22
T48240
hypothetical protein T7H20.240 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48240
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48240
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <BEV>
A;Cross-references: UNIPROT:Q9LZL3; UNIPARC:UPI00000A02B5; EMBL:AL162508
A;Experimental source: cultivar Columbia; BAC clone T7H20
C;Genetics:
A;Map position: 5
A;Note: T7H20.240

Alignment Scores:
Pred. No.: 10.4 Length: 453
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T48240 (1-453)

Qy 1288 TCCTTCAGCAGCAGCTCCAGTTCCCAA 1314
||||||||||||||||||||||||||||
Db 36 SerPheSerSerSerSerSerSerGln 44

RESULT 23
F75524
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: F75524
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: F75524
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <WHI>
A;Cross-references: UNIPROT:Q9RXC3; UNIPARC:UPI00000C1748; GB:AE001899; GB:AE000513; NI
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0392
A;Map position: 1
C;Superfamily: loricrin

Alignment Scores:
Pred. No.: 10.4 Length: 465
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x F75524 (1-465)

Qy 1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
||||||||||||||||||||||||||||
Db 408 SerSerSerSerSerSerGlnSerSer 416

RESULT 24
I38240
transcription factor SOX4 - human
N;Alternate names: sex-determining region Y box 4 protein
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Oct-2004
C;Accession: I38240; S22938; S21479; S31724
R;Farr, C.J.; Easty, D.J.; Ragoussis, J.; Collignon, J.; Lovell-Badge, R.; Goodfellow,
Mamm. Genome 4, 577-584, 1993
A;Title: Characterization and mapping of the human SOX4 gene.
A;Reference number: I38240; MUID:94093204; PMID:8268656
A;Accession: I38240
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-474 <RES>
A;Cross-references: UNIPROT:Q06945; UNIPARC:UPI0000047FA0; EMBL:X70683; NID:g36552; PID
R;Denny, P.; Swift, S.; Brand, N.; Dabhade, N.; Barton, P.; Ashworth, A.
Nucleic Acids Res. 20, 2887, 1992
A;Title: A conserved family of genes related to the testis determining gene, SRY.
A;Reference number: S22935; MUID:92310993; PMID:1614875
A;Accession: S22938
A;Molecule type: mRNA
A;Residues: 70,'P',72-123 <DEN>
A;Cross-references: UNIPARC:UPI000016B04E; EMBL:X65661; NID:g36551; PIDN:CAA46612.1; PI

```
C;Genetics:
A;Gene: GDB:SOX4
A;Cross-references: GDB:250365; OMIM:184430
A;Map position: 6p22.2-6p22.16p23-6p23
C;Keywords: DNA binding; transcription regulation
F;56-131/Domain: HMG box homology <HMG1>

Alignment Scores:
Pred. No.:              10.4          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x I38240 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 25
T37541
probable glycolipid anchored surface protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37541
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21721
A;Accession: T37541
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-510 <MUR>
A;Cross-references: UNIPROT:O13692; UNIPARC:UPI000013AA47; EMBL:Z98595; PIDN:CAB11192.1;
A;Experimental source: strain 972h-; cosmid c11E3
C;Genetics:
A;Gene: SPDB:SPAC11E3.13c
A;Map position: 1
C;Superfamily: glycophospholipid-anchored surface glycoprotein GAS1

Alignment Scores:
Pred. No.:              10.2          Length:        510
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x T37541 (1-510)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        473 SerSerGlySerSerSerGlySerSer 481

RESULT 26
B32372
male-specific doublesex protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 09-Jul-2004
C;Accession: B32372
R;Burtis, K.C.; Baker, B.S.
Cell 56, 997-1010, 1989
A;Title: Drosophila doublesex gene controls somatic sexual differentiation by producing
A;Reference number: A32372; MUID:89168451; PMID:2493994
A;Accession: B32372
A;Status: preliminary
A;Molecule type: DNA; mRNA
A;Residues: 1-549 <BUR>
A;Cross-references: UNIPROT:P23023; UNIPARC:UPI0000007519; GB:M25294; GB:J03156
C;Genetics:
A;Gene: FlyBase:dsx
A;Cross-references: FlyBase:FBgn0000504
C;Keywords: alternative splicing; DNA binding; transcription regulation

Alignment Scores:
Pred. No.:              10.1          Length:        549
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x B32372 (1-549)

Qy        211 TCCAGCAGTTCTGGCAGCTCCATTTCC 237
              ||||||||||||||||||||||||||
Db         29 SerSerSerSerGlySerSerIleSer 37

RESULT 27
S33167
gene pointed protein splice form 1 - fruit fly (Drosophila melanogaster)
N;Alternate names: ETS-like protein
C;Species: Drosophila melanogaster
C;Date: 13-Jan-1995 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S33167; S28819; S28818
R;Klaembt, C.
Development 117, 163-176, 1993
A;Title: The Drosophila gene pointed encodes two ETS-like proteins which are involved i
A;Reference number: S33167; MUID:94038653; PMID:8223245
A;Accession: S33167
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-623 <KLA>
A;Cross-references: UNIPROT:P51022; UNIPARC:UPI000000760C; EMBL:X69166; NID:g288079; PI
R;Chen, T.; Bunting, M.; Karim, F.D.; Thummel, C.S.
Dev. Biol. 151, 176-191, 1992
A;Title: Isolation and characterization of five Drosophila genes that encode an ets-rel
A;Reference number: S28819; MUID:92249640; PMID:1577186
A;Accession: S28819
A;Molecule type: mRNA
A;Residues: 456-613 <CHE>
A;Cross-references: UNIPARC:UPI000016BB75; EMBL:M88472; NID:g157191; PIDN:AAC34200.1; P
R;Pribyl, L.J.; Watson, D.K.; McWilliams, M.J.; Ascione, R.; Papas, T.S.
Dev. Biol. 127, 45-53, 1988
A;Title: The Drosophila ets-2 gene: molecular structure, chromosomal localization, and
A;Reference number: S28818; MUID:88196618; PMID:2834248
A;Accession: S28818
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 445-466,'R',468-603 <PRI>
A;Cross-references: UNIPARC:UPI000016BBE3; EMBL:M20408; NID:g157384; PIDN:AAA28521.1; P
A;Note: the authors translated the codon AGA for residue 467 as Gln
C;Genetics:
A;Gene: FlyBase:pnt
A;Cross-references: FlyBase:FBgn0003118
A;Introns: 550/3
C;Superfamily: pointed protein, splice form 1; ets DNA-binding domain homology
C;Keywords: alternative splicing; DNA binding; nucleus; transcription factor
F;517-595/Domain: ets DNA-binding domain homology <ETS>

Alignment Scores:
Pred. No.:              9.87          Length:        623
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     1             Gaps:          0

US-10-677-865A-2 (1-1590) x S33167 (1-623)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        161 SerGlySerSerSerSerGlySerSer 169

RESULT 28
```

S39356
transcription factor btd - fruit fly (Drosophila sp.)
C;Species: Drosophila sp.
C;Date: 18-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 07-May-1999
C;Accession: S39356
R;Wimmer, E.A.; Jaeckle, H.; Pfeifle, C.; Cohen, S.M.
Nature 366, 690-694, 1993
A;Title: A Drosophila homologue of human Sp1 is a head-specific segmentation gene.
A;Reference number: S39356; MUID:94081952; PMID:8259212
A;Accession: S39356
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-644 <WIM>
A;Cross-references: UNIPARC:UPI0000124C17; EMBL:Z29361; NID:g441283; PID:g441284
C;Genetics:
A;Gene: FlyBase:btd
A;Cross-references: FlyBase:FBgn0000233
A;Introns: 245/2

Alignment Scores:

Pred. No.:	9.81	Length:	644
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x S39356 (1-644)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        109 SerSerGlySerSerSerGlySerSer 117
```

RESULT 29
A82704
1,4-beta-cellobiosidase XF1267 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82704
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82704
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-683 <SIM>
A;Cross-references: UNIPROT:Q9PDW2; UNIPARC:UPI00000C2694; GB:AE003960; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1267

Alignment Scores:

Pred. No.:	9.71	Length:	683
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x A82704 (1-683)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        558 SerSerGlySerSerSerGlySerSer 566
```

RESULT 30
D84428
hypothetical protein At2g01730 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84428
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.,
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84428
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-837 <STO>
A;Cross-references: UNIPROT:Q9ZUA1; UNIPARC:UPI000009F8B2; GB:AE002093; NID:g4220489; P
C;Genetics:
A;Gene: At2g01730
A;Map position: 2

Alignment Scores:

Pred. No.:	9.37	Length:	837
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x D84428 (1-837)

```
Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||||||
Db        600 GlySerPheSerSerSerSerSerSer 608
```

RESULT 31
T13230
dachshund isoform 4 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13230
R;Mardon, G.; Solomon, N.M.; Rubin, G.M.
Development 120, 3473-3486, 1994
A;Title: Dachshund encodes a nuclear protein required for normal eye and leg developmen
A;Reference number: Z17633; MUID:95121203; PMID:7821215
A;Accession: T13230
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1065 <MAR>
A;Cross-references: UNIPROT:Q24029; UNIPARC:UPI0000075247; EMBL:U19269; NID:g624295; PI
C;Genetics:
A;Gene: dachshund
A;Cross-references: FlyBase:FBgn0005677
A;Map position: 2
A;Introns: 312/2
C;Function:
A;Description: required for normal eye and leg development
C;Keywords: alternative splicing; eye; nucleus

Alignment Scores:

Pred. No.:	8.99	Length:	1065
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0

```
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x T13230 (1-1065)

Qy        747 TCAGAGGCCTGTGGTGGTGGTGGTGGA 773
              ||||||||||||||||||||||||||||
Db        839 SerGluAlaCysGlyGlyGlyGlyGly 847

RESULT 32
T13232
dachshund protein isoform 3 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13232
R;Mardon, G.; Solomon, N.M.; Rubin, G.M.
Development 120, 3473-3486, 1994
A;Title: Dachshund encodes a nuclear protein required for normal eye and leg development
A;Reference number: Z17633; MUID:95121203; PMID:7821215
A;Accession: T13232
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1072 <MAR>
A;Cross-references: UNIPROT:Q24031; UNIPARC:UPI000007D9BB; EMBL:U19269; NID:g624295; PID
C;Genetics:
A;Gene: dachshund
A;Cross-references: FlyBase:FBgn0005677
A;Map position: 2
A;Introns: 312/2
C;Function:
A;Description: required for normal eye and leg development
C;Keywords: alternative splicing; eye; nucleus

Alignment Scores:
Pred. No.:              8.98          Length:        1072
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x T13232 (1-1072)

Qy        747 TCAGAGGCCTGTGGTGGTGGTGGTGGA 773
              ||||||||||||||||||||||||||||
Db        846 SerGluAlaCysGlyGlyGlyGlyGly 854

RESULT 33
T13229
dachshund protein isoform 1 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13229; T13228
R;Mardon, G.; Solomon, N.M.; Rubin, G.M.
Development 120, 3473-3486, 1994
A;Title: Dachshund encodes a nuclear protein required for normal eye and leg development
A;Reference number: Z17633; MUID:95121203; PMID:7821215
A;Accession: T13229
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1074 <MAR>
A;Cross-references: UNIPROT:Q24028; UNIPARC:UPI0000080FAD; EMBL:U19269; NID:g624295; PID
A;Accession: T13228
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-320,323-1074 <MA2>
A;Cross-references: UNIPARC:UPI0000082785; EMBL:U19269; NID:g624295; PID:g624298; PIDN:A
C;Genetics:
A;Gene: dachshund
A;Cross-references: FlyBase:FBgn0005677
A;Map position: 2
A;Introns: 321/2
C;Function:
A;Description: is required for normal eye and development
C;Keywords: alternative splicing; eye; nucleus

Alignment Scores:
Pred. No.:              8.97          Length:        1074
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x T13229 (1-1074)

Qy        747 TCAGAGGCCTGTGGTGGTGGTGGTGGA 773
              ||||||||||||||||||||||||||||
Db        848 SerGluAlaCysGlyGlyGlyGlyGly 856

RESULT 34
T13231
dachshund protein isoform 5 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13231
R;Mardon, G.; Solomon, N.M.; Rubin, G.M.
Development 120, 3473-3486, 1994
A;Title: Dachshund encodes a nuclear protein required for normal eye and leg developmen
A;Reference number: Z17633; MUID:95121203; PMID:7821215
A;Accession: T13231
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1081 <MAR>
A;Cross-references: UNIPROT:Q24030; UNIPARC:UPI0000082883; EMBL:U19269; NID:g624295; PI
C;Genetics:
A;Gene: dachshund
A;Cross-references: FlyBase:FBgn0005677
A;Map position: 2
C;Function:
A;Description: required for normal eye and leg development
C;Keywords: alternative splicing; eye; nucleus

Alignment Scores:
Pred. No.:              8.96          Length:        1081
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x T13231 (1-1081)

Qy        747 TCAGAGGCCTGTGGTGGTGGTGGTGGA 773
              ||||||||||||||||||||||||||||
Db        855 SerGluAlaCysGlyGlyGlyGlyGly 863

RESULT 35
B54962
sterol regulatory element binding protein 2 precursor - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;Accession: B54962
R;Yang, J.; Sato, R.; Goldstein, J.L.; Brown, M.S.
Genes Dev. 8, 1910-1919, 1994
A;Title: Sterol-resistant transcription in CHO cells caused by gene rearrangement that
A;Reference number: A54962; MUID:95047343; PMID:7958866
A;Accession: B54962
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1139 <YAN>
A;Cross-references: UNIPROT:Q60429; UNIPARC:UPI0000178926; GB:U12330
A;Note: 493 Ser was also found
C;Superfamily: sterol regulatory element binding protein
C;Keywords: DNA binding; membrane protein

Alignment Scores:
```

Pred. No.: 8.88 Length: 1139
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x B54962 (1-1139)

Qy 253 GGTGGTGGCTCCAGTGGTAGCTCCAGC 279
||||||||||||||||||||||||||||
Db 56 GlyGlyGlySerSerGlySerSerSer 64

RESULT 36
T38949
hypothetical protein SPAC57A7.05 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38949
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21819
A;Accession: T38949
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1337 <SKE>
A;Cross-references: UNIPROT:P87136; UNIPARC:UPI000013AA02; EMBL:Z95396; PIDN:CAB08763.1;
A;Experimental source: strain 972h-; cosmid c57A7
C;Genetics:
A;Gene: SPDB:SPAC57A7.05
A;Map position: 1

Alignment Scores:
Pred. No.: 8.64 Length: 1337
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T38949 (1-1337)

Qy 1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
|||||||||||||||||||||||||||
Db 136 SerSerSerSerSerSerGlnSerSer 144

RESULT 37
T08852
lustrin A - California red abalone
C;Species: Haliotis rufescens (California red abalone)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08852
R;Shen, X.; Belcher, A.M.; Hansma, P.K.; Stucky, G.D.; Morse, D.E.
J. Biol. Chem. 272, 32472-32481, 1997
A;Title: Molecular cloning and characterization of lustrin A, a matrix protein from shel
A;Reference number: Z16496; MUID:98070424; PMID:9405458
A;Accession: T08852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1428 <SHE>
A;Cross-references: UNIPROT:O44341; UNIPARC:UPI00000836D6; EMBL:AF023459; NID:g2723361;
A;Experimental source: tissue type mantle (shell and pearl nacre); cell type pallial
C;Keywords: extracellular matrix; extracellular protein
F;1382-1426/Domain: antileukoproteinase repeat homology <ALP>

Alignment Scores:
Pred. No.: 8.54 Length: 1428
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T08852 (1-1428)

Qy 262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
||||||||||||||||||||||||||||
Db 1029 SerSerGlySerSerSerGlySerSer 1037

RESULT 38
T09219
basal transcription factor SNAPc large chain SNAP190 - human
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 05-Oct-2004
C;Accession: T09219
R;Wong, M.W.; Henry, R.W.; Ma, B.; Kobayashi, R.; Klages, N.; Matthias, P.; Strubin, M.
Mol. Cell. Biol. 18, 368-377, 1998
A;Title: The large subunit of basal transcription factor SNAPc is a Myb domain protein
A;Reference number: Z16616; MUID:98078693; PMID:9418884
A;Accession: T09219
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1469 <WON>
A;Cross-references: UNIPROT:Q9Y6P7; UNIPARC:UPI0000073E1B; EMBL:AF032387; NID:g2641556;
A;Experimental source: tissue type fetal cell teratocarcinoma
C;Genetics:
A;Gene: SNAP190
C;Function:
A;Description: transcription factor; required for transcription of snRNA genes
C;Keywords: DNA binding; RNA biosynthesis; transcription factor
F;396-447/Domain: myb DNA-binding repeat homology <MYB>

Alignment Scores:
Pred. No.: 8.5 Length: 1469
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T09219 (1-1469)

Qy 262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
||||||||||||||||||||||||||||
Db 523 SerSerGlySerSerSerGlySerSer 531

RESULT 39
A45546
major merozoite surface antigen precursor - Plasmodium chabaudi chabaudi
C;Species: Plasmodium chabaudi chabaudi
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004
C;Accession: A45546
R;Deleersnijder, W.; Hendrix, D.; Bendahman, N.; Hanegreefs, J.; Brijs, L.; Hamers-Cast
Mol. Biochem. Parasitol. 43, 231-244, 1990
A;Title: Molecular cloning and sequence analysis of the gene encoding the major merozoi
A;Reference number: A45546; MUID:91218805; PMID:2090945
A;Accession: A45546
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1785 <DEL>
A;Cross-references: UNIPROT:Q25685; UNIPARC:UPI000007B6DB; GB:M34947; NID:g160597; PID:
C;Superfamily: G surface protein
C;Keywords: surface antigen

Alignment Scores:
Pred. No.: 8.21 Length: 1785
Score: 9.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.7% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x A45546 (1-1785)

Qy 254 GTGGTGGCTCCAGTGGTAGCTCCAGCG 280
|||||||||||||||||||||||||||

```
Db       1341 ValValAlaProValValAlaProAla 1349

RESULT 40
S75251
hypothetical protein slr1028 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S75251
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75251
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-3972 <KAN>
A;Cross-references: UNIPROT:P73139; UNIPARC:UPI00000C0CAB; EMBL:D90904; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Synechocystis hypothetical protein slr1028

Alignment Scores:
Pred. No.:                7.14          Length:        3972
Score:                    9.00          Matches:       9
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.7%          Indels:        0
DB:                       2             Gaps:          0

US-10-677-865A-2 (1-1590) x S75251 (1-3972)

Qy        199 AGCTACAGTGGCTCCAGCAGTTCTGGC 225
              ||||||||||||||||||||||||||
Db        752 SerTyrSerGlySerSerSerSerGly 760

RESULT 41
T45015
hypothetical protein [imported] - Rhodobacter sphaeroides
C;Species: Rhodobacter sphaeroides
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45015
R;Hamblin, P.A.; Maguire, B.A.; Grishanin, R.N.; Armitage, J.P.
Mol. Microbiol. 26, 1083-1096, 1997
A;Title: Evidence for dual chemosensory pathways in Rhodobacter sphaeroides.
A;Reference number: Z22898; MUID:98086113; PMID:9426144
A;Accession: T45015
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-89 <HAM>
A;Cross-references: UNIPROT:O33552; UNIPARC:UPI00000B0BE5; EMBL:AJ000977; PIDN:CAA04427.
A;Experimental source: strain WS8
C;Genetics:
A;Note: orf10

Alignment Scores:
Pred. No.:                126           Length:        89
Score:                    8.00          Matches:       8
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.5%          Indels:        0
DB:                       2             Gaps:          0

US-10-677-865A-2 (1-1590) x T45015 (1-89)

Qy        228 GCTGCCAGAACTGCTGGAGCCACT 205
              ||||||||||||||||||||||||
Db         52 AlaAlaArgThrAlaGlyAlaThr 59

RESULT 42
ERBP1Z
gene 1C protein - phage PZA
C;Species: phage PZA
A;Note: host Bacillus subtilis
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 05-Oct-2004
C;Accession: C24528
R;Paces, V.; Vlcek, C.; Urbanek, P.; Hostomsky, Z.
Gene 38, 45-56, 1985
A;Title: Nucleotide sequence of the major early region of Bacillus subtilis phage PZA,
A;Reference number: A91538; MUID:86056991; PMID:3934048
A;Accession: C24528
A;Molecule type: DNA
A;Residues: 1-97 <PAC>
A;Cross-references: UNIPROT:P06949; UNIPARC:UPI0000138508; GB:M11813; GB:M13904; GB:M13
C;Genetics:
A;Gene: 1C
C;Superfamily: early protein GP1C
C;Keywords: early protein

Alignment Scores:
Pred. No.:                124           Length:        97
Score:                    8.00          Matches:       8
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.5%          Indels:        0
DB:                       1             Gaps:          0

US-10-677-865A-2 (1-1590) x ERBP1Z (1-97)

Qy        984 GTTCTCTTTGGTGAAGTAGCCCAC 961
              ||||||||||||||||||||||||
Db         87 ValLeuPheGlyGluValAlaHis 94

RESULT 43
D86931
hypothetical protein ML0180 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: D86931
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: D86931
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <STO>
A;Cross-references: UNIPROT:Q9CD66; UNIPARC:UPI00000C6C60; GB:AL450380; NID:g13092539;
C;Genetics:
A;Gene: ML0180

Alignment Scores:
Pred. No.:                123           Length:        99
Score:                    8.00          Matches:       8
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.5%          Indels:        0
DB:                       2             Gaps:          0

US-10-677-865A-2 (1-1590) x D86931 (1-99)

Qy        265 AGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||
Db         76 SerGlySerSerSerGlySerSer 83

RESULT 44
A72572
hypothetical protein APE1858 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: A72572
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
```

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: A72572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-100 <KAW>
A;Cross-references: UNIPROT:Q9YAT6; UNIPARC:UPI000005E0FA; DDBJ:AP000062; NID:g5105244;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1858
C;Superfamily: Aeropyrum pernix hypothetical protein APE1858

Alignment Scores:
Pred. No.: 123 Length: 100
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x A72572 (1-100)

Qy 214 AGCAGTTCTGGCAGCTCCATTTCC 237
||||||||||||||||||||||||
Db 47 SerSerSerGlySerSerIleSer 54

RESULT 45
D61615
sericin MG-1 - greater wax moth (fragment)
N;Alternate names: silk protein MG-1
C;Species: Galleria mellonella (greater wax moth)
C;Date: 18-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004
C;Accession: D61615
R;Zurovec, M.; Sehnal, F.; Scheller, K.; Kumaran, A.K.
Insect Biochem. Mol. Biol. 22, 55-67, 1992
A;Title: Silk gland specific cDNAs from Galleria mellonella L.
A;Reference number: A61615
A;Accession: D61615
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-115 <ZUR>
A;Cross-references: UNIPROT:O96614; UNIPARC:UPI0000135866
C;Keywords: tandem repeat

Alignment Scores:
Pred. No.: 120 Length: 115
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x D61615 (1-115)

Qy 265 AGTGGTAGCTCCAGCGGATCCAGC 288
||||||||||||||||||||||||
Db 92 SerGlySerSerSerGlySerSer 99

RESULT 46
T45451
hypothetical protein MLCB373.32c [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45451
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z22967
A;Accession: T45451
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-120 <JAM>
A;Cross-references: UNIPROT:Q9Z5G3; UNIPARC:UPI00000BAC75; EMBL:AL035500; PIDN:CAB36694
A;Experimental source: cosmid L373
C;Genetics:
A;Note: MLCB373.32c

Alignment Scores:
Pred. No.: 119 Length: 120
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T45451 (1-120)

Qy 265 AGTGGTAGCTCCAGCGGATCCAGC 288
||||||||||||||||||||||||
Db 97 SerGlySerSerSerGlySerSer 104

RESULT 47
A45519
variant surface glycoprotein pTb1.5-c2 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 09-Jul-2004
C;Accession: A45519
R;Smiley, B.L.; Aline Jr., R.F.; Myler, P.J.; Stuart, K.
Mol. Biochem. Parasitol. 42, 143-152, 1990
A;Title: A retroposon in the 5' flank of a Trypanosoma brucei VSG gene lacks insertiona
A;Reference number: A45519; MUID:91042830; PMID:2172817
A;Accession: A45519
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-120 <SMI>
A;Cross-references: UNIPROT:Q26834; UNIPARC:UPI0000078E35; GB:M33486; NID:g162456; PID:
C;Keywords: glycoprotein

Alignment Scores:
Pred. No.: 119 Length: 120
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x A45519 (1-120)

Qy 407 GCAGCAGCTCTCACTCGGGAAGCA 430
||||||||||||||||||||||||
Db 11 AlaAlaAlaLeuThrArgGluAla 18

RESULT 48
S69871
hypothetical protein YML012c-a - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 23-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C;Accession: S69871
R;Gentles, S.; Bowman, S.
submitted to the EMBL Data Library, June 1995
A;Reference number: S55102
A;Accession: S69871
A;Molecule type: DNA
A;Residues: 1-125 <GEN>
A;Cross-references: UNIPARC:UPI00001683C5; EMBL:Z49810; GSPDB:GN00013; MIPS:YML012c-a
C;Genetics:
A;Gene: MIPS:YML012c-a
A;Map position: 13L

Alignment Scores:
Pred. No.: 118 Length: 125
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0

```
DB:                    2             Gaps:          0

US-10-677-865A-2 (1-1590) x S69871 (1-125)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db          8 SerPheSerSerSerSerSerSer 15

RESULT 49
T00282
hypothetical protein - Escherichia coli plasmid pO157
C;Species: Escherichia coli
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00282
R;Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,
 S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A;Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrha
A;Reference number: Z14127; MUID:98290540; PMID:9628576
A;Accession: T00282
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-129 <MAK>
A;Cross-references: UNIPROT:O82903; UNIPARC:UPI00000B75B1; EMBL:AB011549; NID:g4589740;
A;Experimental source: strain EHEC O157:H7, substrain RIMD 0509952
C;Genetics:
A;Genome: plasmid pO157

Alignment Scores:
Pred. No.:             118           Length:        129
Score:                 8.00          Matches:       8
Percent Similarity:    100.0%        Conservative:  0
Best Local Similarity: 100.0%        Mismatches:    0
Query Match:           1.5%          Indels:        0
DB:                    2             Gaps:          0

US-10-677-865A-2 (1-1590) x T00282 (1-129)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         23 SerPheSerSerSerSerSerSer 30

RESULT 50
A86417
probable auxin-induced protein, 45653-45228 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C;Accession: A86417
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A86417
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-141 <STO>
A;Cross-references: UNIPROT:Q9C7Q8; UNIPARC:UPI00000A8C61; GB:AE005172; NID:g10092226; P
C;Genetics:
A;Map position: 1
C;Superfamily: auxin-induced protein 10A

Alignment Scores:
Pred. No.:             116           Length:        141
Score:                 8.00          Matches:       8
Percent Similarity:    100.0%        Conservative:  0
Best Local Similarity: 100.0%        Mismatches:    0
Query Match:           1.5%          Indels:        0
DB:                    2             Gaps:          0

US-10-677-865A-2 (1-1590) x A86417 (1-141)

Qy        229 TCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||
Db        118 SerIleSerSerAlaArgSerSer 125

RESULT 51
H72610
hypothetical protein APE1348 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H72610
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-143 <KAW>
A;Cross-references: UNIPROT:Q9YCA6; UNIPARC:UPI000005DEF3; DDBJ:AP000061; NID:g5104821;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1348

Alignment Scores:
Pred. No.:             116           Length:        143
Score:                 8.00          Matches:       8
Percent Similarity:    100.0%        Conservative:  0
Best Local Similarity: 100.0%        Mismatches:    0
Query Match:           1.5%          Indels:        0
DB:                    2             Gaps:          0

US-10-677-865A-2 (1-1590) x H72610 (1-143)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGC 228
              ||||||||||||||||||||||||
Db         27 SerGlySerSerSerSerGlySer 34

RESULT 52
JC5398
leukotriene C4 synthase (EC 6.-.-.-) - human
C;Species: Homo sapiens (man)
C;Date: 20-Jun-1997 #sequence_revision 20-Jun-1997 #text_change 20-Jun-2000
C;Accession: JC5398
R;Surapureddi, S.; Morgenstern, R.; Soederstroem, M.; Hammarstroem, S.
Biochem. Biophys. Res. Commun. 229, 388-395, 1996
A;Title: Interaction of human leukotriene C4 synthase and microsomal glutathione transf
A;Reference number: JC5398; MUID:97127413; PMID:8954909
A;Accession: JC5398
A;Molecule type: mRNA
A;Residues: 1-150 <SUR>
A;Cross-references: UNIPARC:UPI00001796CE
A;Experimental source: lung
C;Comment: This enzyme binds to microsomal glutathione transferase in the presence of M
C;Superfamily: Synechocystis hypothetical protein sll1147
C;Keywords: ligase

Alignment Scores:
Pred. No.:             115           Length:        150
Score:                 8.00          Matches:       8
Percent Similarity:    100.0%        Conservative:  0
Best Local Similarity: 100.0%        Mismatches:    0
Query Match:           1.5%          Indels:        0
DB:                    2             Gaps:          0

US-10-677-865A-2 (1-1590) x JC5398 (1-150)

Qy        137 CGCGTATCACCTCCCCTAACGACC 160
```

```
                    ||||||||||||||||||||||
Db          34 ArgValSerProProLeuThrThr 41

RESULT 53
I38595
leukotriene-C4 synthase (EC 2.5.1.37) - human
N;Alternate names: leukotriene C4 synthase (glutathione S-transferase)
C;Species: Homo sapiens (man)
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 09-Jul-2004
C;Accession: I38595; A47263; I38705
R;Lam, B.K.; Penrose, J.F.; Freeman, G.J.; Austen, K.F.
Proc. Natl. Acad. Sci. U.S.A. 91, 7663-7667, 1994
A;Title: Expression cloning of a cDNA for human leukotriene C4 synthase, an integral mem
A;Reference number: I38595; MUID:94329576; PMID:8052639
A;Accession: I38595
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <RES>
A;Cross-references: UNIPROT:Q16873; UNIPARC:UPI0000043F83; EMBL:U09353; NID:g520484; PID
R;Nicholson, D.W.; Ali, A.; Vaillancourt, J.P.; Calaycay, J.R.; Mumford, R.A.; Zamboni,
Proc. Natl. Acad. Sci. U.S.A. 90, 2015-2019, 1993
A;Title: Purification to homogeneity and the N-terminal sequence of human leukotriene C4
A;Reference number: A47263; MUID:93189628; PMID:8446623
A;Accession: A47263
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20,'G',22-30,'I',32-35 <NIC>
A;Cross-references: UNIPARC:UPI000014F9AF
A;Experimental source: monocytic leukemia cell line THP-1
A;Note: sequence extracted from NCBI backbone (NCBIP:126925)
R;Welsch, D.J.; Creely, D.P.; Hauser, S.D.; Mathis, K.J.; Krivi, G.G.; Isakson, P.I.
Proc. Natl. Acad. Sci. U.S.A. 91, 9745-9749, 1994
A;Title: Molecular cloning and expression of human leukotriene-C4 synthase.
A;Reference number: I38705; MUID:95024037; PMID:7937884
A;Accession: I38705
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <WEL>
A;Cross-references: UNIPARC:UPI0000043F83; EMBL:U11552; NID:g508820; PIDN:AAA50555.1; PI
C;Genetics:
A;Gene: GDB:LTC4S
A;Cross-references: GDB:384080; OMIM:246530
A;Map position: 5q35-5q35
C;Superfamily: Synechocystis hypothetical protein sll1147
C;Keywords: transferase; transmembrane protein

Alignment Scores:
Pred. No.:                115             Length:         150
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x I38595 (1-150)

Qy        137 CGCGTATCACCTCCCCTAACGACC 160
                ||||||||||||||||||||||
Db         34 ArgValSerProProLeuThrThr 41

RESULT 54
S56258
hypothetical protein YFR003c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S56258
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces ce
A;Reference number: S56186
A;Accession: S56258
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-155 <MUR>
A;Cross-references: UNIPROT:P43587; UNIPARC:UPI000013AE86; EMBL:D50617; NID:g836685; PI
C;Genetics:
A;Cross-references: SGD:S0001899
A;Map position: 6R

Alignment Scores:
Pred. No.:                114             Length:         155
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x S56258 (1-155)

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCC 231
                ||||||||||||||||||||||||
Db         92 GlySerSerSerSerGlySerSer 99

RESULT 55
A36107
ribosomal protein L35 precursor, chloroplast - spinach
C;Species: Spinacia oleracea (spinach)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A36107
R;Smooker, P.M.; Choli, T.; Subramanian, A.R.
Biochemistry 29, 9733-9736, 1990
A;Title: Ribosomal protein L35: identification in spinach chloroplasts and isolation of
A;Reference number: A36107; MUID:91104876; PMID:2271612
A;Accession: A36107
A;Molecule type: mRNA
A;Residues: 1-159 <SMO>
A;Cross-references: UNIPROT:P23326; UNIPARC:UPI0000133AD7; GB:J02928
C;Superfamily: Escherichia coli ribosomal protein L35
C;Keywords: chloroplast; protein biosynthesis; ribosome

Alignment Scores:
Pred. No.:                113             Length:         159
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x A36107 (1-159)

Qy       1198 AGCAGTTCTAGCATTTCCAGCAGC 1221
                ||||||||||||||||||||||||
Db         49 SerSerSerSerIleSerSerSer 56

RESULT 56
B96842
hypothetical protein F23A5.28 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B96842
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B96842
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-163 <STO>
A;Cross-references: UNIPROT:Q9SAG8; UNIPARC:UPI00000A712D; GB:AE005173; NID:g6503304; P
```

C;Genetics:
A;Gene: F23A5.28
A;Map position: 1

Alignment Scores:
Pred. No.: 113 Length: 163
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x B96842 (1-163)

Qy 1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
||||||||||||||||||||||||
Db 19 SerPheSerSerSerSerSerSer 26

RESULT 57
T14187
hypothetical protein T28D5.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14187
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z17931
A;Accession: T14187
A;Molecule type: DNA
A;Residues: 1-164 <BEV>
A;Cross-references: UNIPROT:Q9STN5; UNIPARC:UPI000000C52B; EMBL:AL109819
A;Experimental source: cultivar Columbia; BAC clone T28D5
C;Genetics:
A;Gene: ATSP:T28D5.20
A;Map position: 4
A;Introns: 39/2
C;Superfamily: Arabidopsis thaliana hypothetical protein T28D5.20

Alignment Scores:
Pred. No.: 113 Length: 164
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x T14187 (1-164)

Qy 1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
||||||||||||||||||||||||
Db 23 SerPheSerSerSerSerSerSer 30

RESULT 58
A03864
hypothetical 17.7K protein - human adenovirus 2
C;Species: Mastadenovirus h2 (human adenovirus 2)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: F92351; F92352; A03864
R;Gingeras, T.R.; Sciaky, D.; Gelinas, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull
J. Biol. Chem. 257, 13475-13491, 1982
A;Title: Nucleotide sequences from the adenovirus-2 genome.
A;Reference number: A92351; MUID:83056843; PMID:7142161
A;Accession: F92351
A;Molecule type: DNA
A;Residues: 1-168 <GIN>
A;Cross-references: UNIPROT:P03292; UNIPARC:UPI00001393BA
R;Alestrom, P.; Akusjarvi, G.; Pettersson, M.; Pettersson, U.
J. Biol. Chem. 257, 13492-13498, 1982
A;Title: DNA sequence analysis of the region encoding the terminal protein and the hypot
A;Reference number: A92352; MUID:83056844; PMID:7142162
A;Accession: F92352
A;Molecule type: DNA
A;Residues: 1-168 <ALE>
A;Cross-references: UNIPARC:UPI00001393BA
C;Superfamily: human adenovirus 2 hypothetical 17.7K protein

Alignment Scores:
Pred. No.: 112 Length: 168
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x A03864 (1-168)

Qy 1198 AGCAGTTCTAGCATTTCCAGCAGC 1221
||||||||||||||||||||||||
Db 18 SerSerSerSerIleSerSerSer 25

RESULT 59
S41755
cyclin E type II - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C;Accession: S41755
R;Richardson, H.E.; O'Keefe, L.V.; Reed, S.I.; Saint, R.
Development 119, 673-690, 1993
A;Title: A Drosophila G(1)-specific cyclin E homolog exhibits different modes of expres
A;Reference number: S41755; MUID:94244469; PMID:8187637
A;Accession: S41755
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-173 <RIC>
A;Cross-references: UNIPARC:UPI000016BD2C; EMBL:X75027; NID:g429167; PIDN:CAA52935.1; P
C;Genetics:
A;Gene: FlyBase:CycE
A;Cross-references: FlyBase:FBgn0010382
C;Keywords: cell cycle control

Alignment Scores:
Pred. No.: 112 Length: 173
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x S41755 (1-173)

Qy 1198 AGCAGTTCTAGCATTTCCAGCAGC 1221
||||||||||||||||||||||||
Db 39 SerSerSerSerIleSerSerSer 46

RESULT 60
PN0109
keratin-like protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
C;Accession: PN0109
R;Shvets, Y.P.; Chumakov, I.M.; Kisselev, L.L.
Mol. Biol. 24, 663-677, 1990
A;Title: Sequencing the fragment from a transcriptionally active rat genomic locus and
A;Reference number: PN0109
A;Accession: PN0109
A;Molecule type: mRNA
A;Residues: 1-183 <SHV>
A;Cross-references: UNIPARC:UPI0000177375
C;Genetics:
A;Gene: K51
C;Superfamily: loricrin
C;Keywords: duplication
F;21-28/Region: 8-residue repeat
F;34-41/Region: 8-residue repeat
F;73-88/Region: 16-residue repeat
F;77-84/Region: 8-residue repeat

F;89-104/Region: 16-residue repeat
F;93-100/Region: 8-residue repeat
F;105-112/Region: 8-residue repeat
F;118-125/Region: 8-residue repeat
F;152-163/Region: 12-residue repeat
F;152-159/Region: 8-residue repeat
F;164-175/Region: 12-residue repeat
F;164-171/Region: 8-residue repeat
F;176-183/Region: 8-residue repeat

Alignment Scores:

Pred. No.:	111	Length:	183
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x PN0109 (1-183)

```
Qy        247 AGCTCTGGTGGTGGCTCCAGTGGT 270
              ||||||||||||||||||||||||
Db         52 SerSerGlyGlyGlySerSerGly 59
```

RESULT 61
F71250
probable adenine phosphoribosyltransferase (apt) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: F71250
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: F71250
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-190 <COL>
A;Cross-references: UNIPROT:O84001; UNIPARC:UPI0000125CF4; GB:AE001271; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP1039
C;Superfamily: adenine phosphoribosyltransferase

Alignment Scores:

Pred. No.:	110	Length:	190
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x F71250 (1-190)

```
Qy        825 ATTGCTACAGGGGGGACCTTGAAC 802
              ||||||||||||||||||||||||
Db        135 IleAlaThrGlyGlyThrLeuAsn 142
```

RESULT 62
T01425
hypothetical protein T2H3.15 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01425
R;Kalicki, J.; Gibson, A.
submitted to the EMBL Data Library, August 1998
A;Description: The sequence of A. thaliana T2H3.
A;Reference number: Z14324
A;Accession: T01425
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-196 <KAL>
A;Cross-references: UNIPROT:O81421; UNIPARC:UPI00000A1E39; EMBL:AF075597; NID:g3298610;
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 77/1
A;Note: T2H3.15

Alignment Scores:

Pred. No.:	109	Length:	196
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T01425 (1-196)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db          5 SerPheSerSerSerSerSerSer 12
```

RESULT 63
S55621
hypothetical protein 27 - equine herpesvirus 2
C;Species: equine herpesvirus 2
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55621
R;Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A;Title: The DNA sequence of equine herpesvirus 2.
A;Reference number: S55594; MUID:95302501; PMID:7783207
A;Accession: S55621
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-197 <TEL>
A;Cross-references: UNIPROT:Q66630; UNIPARC:UPI00000EF0BD; GB:U20824; NID:g695172; PIDN
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C;Superfamily: equine herpesvirus 2 hypothetical protein 27

Alignment Scores:

Pred. No.:	109	Length:	197
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x S55621 (1-197)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         61 SerPheSerSerSerSerSerSer 68
```

RESULT 64
S60449
probable membrane protein YGR160w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G7004
C;Species: Saccharomyces cerevisiae
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C;Accession: S60449; S64469; S64471
R;Skala, J.; Nawrocki, A.; Goffeau, A.
Yeast 11, 1421-1427, 1995
A;Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Saccha
.
A;Reference number: S60435; MUID:96158062; PMID:8585325
A;Accession: S60449
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <SKA>
A;Cross-references: UNIPROT:P53288; UNIPARC:UPI000013AF87; EMBL:X85807; NID:g1045249; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995
R;Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; De

```
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64428
A;Accession: S64469
A;Molecule type: DNA
A;Residues: 1-203 <VAN>
A;Cross-references: UNIPARC:UPI000013AF87; EMBL:Z72944; NID:g1323270; PID:e243543; PID:g
A;Experimental source: strain S288C
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64471
A;Molecule type: DNA
A;Residues: 8-203 <RIE>
A;Cross-references: UNIPARC:UPI0000168F39; EMBL:Z72944; MIPS:YGR160w
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:FYV13
A;Cross-references: SGD:S0003392
A;Map position: 7R
C;Keywords: transmembrane protein
F;181-197/Domain: transmembrane #status predicted <TMM>

Alignment Scores:
Pred. No.:                109             Length:         203
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x S60449 (1-203)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         90 SerPheSerSerSerSerSerSer 97

RESULT 65
T48207
probable zinc finger protein - Arabidopsis thaliana
N;Alternate names: protein T20L15.130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48207
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48207
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <BEV>
A;Cross-references: UNIPROT:Q9LZW0; UNIPARC:UPI000009F763; EMBL:AL162351
A;Experimental source: cultivar Columbia; BAC clone T20L15
C;Genetics:
A;Map position: 5
A;Note: T20L15.130

Alignment Scores:
Pred. No.:                108             Length:         215
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x T48207 (1-215)

Qy       1198 AGCAGTTCTAGCATTTCCAGCAGC 1221
              ||||||||||||||||||||||||
Db         12 SerSerSerSerIleSerSerSer 19

RESULT 66
AE2461
hypothetical protein all5245 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2461
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2461
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-223 <KUR>
A;Cross-references: UNIPROT:Q8YLQ1; UNIPARC:UPI00000CEDF2; GB:BA000019; PIDN:BAB76944.1
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all5245

Alignment Scores:
Pred. No.:                107             Length:         223
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x AE2461 (1-223)

Qy        582 AAGCTCTTCCTCTTCCCAAACCTC 605
              |||||||||||||||||||||||
Db         25 LysLeuPheLeuPheProAsnLeu 32

RESULT 67
T26532
hypothetical protein Y18D10A.12 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26532
R;Harris, B.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z20226
A;Accession: T26532
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-228 <WIL>
A;Cross-references: UNIPROT:Q9XW11; UNIPARC:UPI0000081F00; EMBL:AL034393; PIDN:CAA22323
A;Experimental source: clone Y18D10A
C;Genetics:
A;Gene: CESP:Y18D10A.12
A;Introns: 83/3; 119/1

Alignment Scores:
Pred. No.:                106             Length:         228
Score:                    8.00            Matches:        8
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.5%            Indels:         0
DB:                       2               Gaps:           0

US-10-677-865A-2 (1-1590) x T26532 (1-228)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         32 SerPheSerSerSerSerSerSer 39

RESULT 68
S41043
rum1 protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 09-Jul-2004
C;Accession: S41043
R;Moreno, S.; Nurse, P.
Nature 367, 236-242, 1994
```

A;Title: Regulation of progression through the G1 phase of the cell cycle by the rum1(+)
A;Reference number: S41043; MUID:94166876; PMID:8121488
A;Accession: S41043
A;Molecule type: DNA
A;Residues: 1-230 <MOR>
A;Cross-references: UNIPROT:P40380; UNIPARC:UPI00001691A7; EMBL:X77730; NID:g456668; PID
C;Genetics:
A;Gene: rum1
A;Map position: 2
C;Superfamily: Schizosaccharomyces rum1 protein

Alignment Scores:

Pred. No.:	106	Length:	230
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x S41043 (1-230)

```
Qy       1186 TCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||
Db        188 SerArgValProSerSerSerSer 195
```

RESULT 69
T40233
Rum1p - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40233
R;Moreno, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21915
A;Accession: T40233
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-230 <MOR>
A;Cross-references: UNIPROT:P40380; UNIPARC:UPI000013531D; EMBL:AL023796; PIDN:CAA19370.
A;Experimental source: strain 972h-; cosmid c32F12
C;Genetics:
A;Gene: SPDB:SPBC32F12.09
A;Map position: 2
C;Superfamily: Schizosaccharomyces rum1 protein

Alignment Scores:

Pred. No.:	106	Length:	230
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T40233 (1-230)

```
Qy       1186 TCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||
Db        188 SerArgValProSerSerSerSer 195
```

RESULT 70
T22915
hypothetical protein F58E10.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22915
R;McMurray, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19637
A;Accession: T22915
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-242 <WIL>
A;Cross-references: UNIPROT:Q9XUW7; UNIPARC:UPI000007B001; EMBL:Z81555; PIDN:CAB04516.1;
A;Experimental source: clone F58E10
C;Genetics:
A;Gene: CESP:F58E10.5
A;Map position: 5
A;Introns: 86/1; 116/1; 201/2

Alignment Scores:

Pred. No.:	105	Length:	242
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x T22915 (1-242)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db          5 SerPheSerSerSerSerSerSer 12
```

RESULT 71
I62393
outer membrane protein ompA - Escherichia vulneris (ATCC 33822) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia vulneris
A;Variety: ATCC 33822
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62393
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62393
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-244 <RES>
A;Cross-references: UNIPROT:Q47880; UNIPARC:UPI00000BBE9D; GB:M63350; NID:g147000; PIDN
A;Experimental source: ATCC 33822
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;104-115/Region: alanine/proline-rich

Alignment Scores:

Pred. No.:	105	Length:	244
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x I62393 (1-244)

```
Qy        257 GTGGCTCCAGTGGTAGCTCCAGCG 280
              ||||||||||||||||||||||||
Db        103 ValAlaProValValAlaProAla 110
```

RESULT 72
I62389
outer membrane protein ompA - Escherichia vulneris (ATCC 33821) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia vulneris
A;Variety: ATCC 33821
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62389
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62389
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

A;Residues: 1-244 <RES>
A;Cross-references: UNIPROT:Q99115; UNIPARC:UPI00000BBF1D; GB:M63349; NID:g146992; PIDN:
A;Experimental source: ATCC 33821
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;104-115/Region: alanine/proline-rich

Alignment Scores:

Pred. No.:	105	Length:	244
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x I62389 (1-244)

```
Qy        257 GTGGCTCCAGTGGTAGCTCCAGCG 280
              ||||||||||||||||||||||||
Db        103 ValAlaProValValAlaProAla 110
```

RESULT 73
A96806
hypothetical protein T5M16.23 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96806
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-244 <STO>
A;Cross-references: UNIPROT:Q9CAP4; UNIPARC:UPI00000A4189; GB:AE005173; NID:g6382508; PI
C;Genetics:
A;Gene: T5M16.23
A;Map position: 1

Alignment Scores:

Pred. No.:	105	Length:	244
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x A96806 (1-244)

```
Qy        457 TCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||
Db         18 SerHisSerSerSerSerSerSer 25
```

RESULT 74
H72314
hypothetical protein TM0928 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: H72314; S44349
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: H72314
A;Molecule type: DNA
A;Residues: 1-246 <ARN>
A;Cross-references: UNIPROT:P56728; UNIPARC:UPI0000139E2A; GB:AE001757; GB:AE000512; NI
A;Experimental source: strain MSB8
R;Darimont, B.; Sterner, R.
EMBO J. 13, 1772-1781, 1994
A;Title: Sequence, assembly and evolution of a primordial ferredoxin from Thermotoga ma
A;Reference number: S44349; MUID:94222023; PMID:8168477
A;Accession: S44349
A;Molecule type: DNA
A;Residues: 201-228,'ERFYGKIEKNV' <DAR>
A;Cross-references: UNIPARC:UPI0000179BB6
C;Genetics:
A;Gene: TM0928
C;Superfamily: Thermotoga maritima hypothetical protein TM0928

Alignment Scores:

Pred. No.:	105	Length:	246
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x H72314 (1-246)

```
Qy       1310 GAACTGGAGCTGCTGCTGAAGGAG 1287
              ||||||||||||||||||||||||
Db          4 GluLeuGluLeuLeuLeuLysGlu 11
```

RESULT 75
C96779
unknown protein F9E10.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C96779
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: UNIPROT:Q9S7M2; UNIPARC:UPI00000A3C3D; GB:AE005173; NID:g6646769; P
C;Genetics:
A;Gene: F9E10.20
A;Map position: 1

Alignment Scores:

Pred. No.:	105	Length:	249
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x C96779 (1-249)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         88 SerPheSerSerSerSerSerSer 95
```

Search completed: March 21, 2006, 22:28:04
Job time : 67 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:15:04 ; Search time 42 Seconds
(without alignments)
5341.857 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4332514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152104_26789/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=oligo.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=50
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_192_@runat_21032006_152104_26789 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	245	46.4	529	1	CDSN_HUMAN	Q15517 homo sapien
2	205	38.8	529	1	CDSN_PANTR	Q7yr44 pan troglod
3	95	18.0	133	2	Q7Z560_HUMAN	Q7z560 homo sapien
4	73	13.8	133	2	Q7Z559_HUMAN	Q7z559 homo sapien
5	55	10.4	534	1	CDSN_MACMU	Q5tm45 macaca mula
6	23	4.4	225	1	CDSN_PIG	O19084 sus scrofa
7	18	3.4	561	1	CDSN_MOUSE	Q7tpc1 mus musculu
8	12	2.3	262	2	Q767M8_PIG	Q767m8 sus scrofa
9	12	2.3	1317	1	GAP_CAEEL	P34288 caenorhabdi
10	12	2.3	1344	2	Q60YS0_CAEBR	Q60ys0 caenorhabdi
11	11	2.1	208	2	Q7PFN9_ANOGA	Q7pfn9 anopheles g
12	11	2.1	262	2	Q8MYJ1_DICDI	Q8myj1 dictyosteli
13	11	2.1	698	2	Q7XJL9_ARATH	Q7xjl9 arabidopsis
14	11	2.1	708	2	Q8SX98_DROME	Q8sx98 drosophila
15	11	2.1	709	1	TMCC2_HUMAN	O75069 homo sapien
16	11	2.1	751	2	Q8MLU0_DROME	Q8mlu0 drosophila
17	11	2.1	761	2	Q7KVL6_DROME	Q7kvl6 drosophila
18	11	2.1	2103	2	Q5K4K9_CHLRE	Q5k4k9 chlamydomon
19	10	1.9	98	2	Q6MAJ1_PARUW	Q6maj1 parachlamyd
20	10	1.9	129	2	Q6MG23_RAT	Q6mg23 rattus norv
21	10	1.9	158	2	Q4RZX2_TETNG	Q4rzx2 tetraodon n
22	10	1.9	173	2	Q569P6_BRARE	Q569p6 brachydanio
23	10	1.9	206	2	Q21636_CAEEL	Q21636 caenorhabdi
24	10	1.9	364	2	Q9R3S1_ACTPL	Q9r3s1 actinobacil
25	10	1.9	365	1	OMLA_ACTPL	Q02937 actinobacil
26	10	1.9	365	2	Q9R3L8_ACTPL	Q9r3l8 actinobacil
27	10	1.9	365	2	Q9RCG1_ACTPL	Q9rcg1 actinobacil
28	10	1.9	368	2	Q9RIH8_ACTPL	Q9rih8 actinobacil
29	10	1.9	368	2	Q9RCG3_ACTPL	Q9rcg3 actinobacil
30	10	1.9	374	1	YB1E_SCHPO	P87179 schizosacch
31	10	1.9	375	2	Q9RCG2_ACTPL	Q9rcg2 actinobacil
32	10	1.9	375	2	Q9A8B9_CAUCR	Q9a8b9 caulobacter
33	10	1.9	406	2	Q8NAC5_HUMAN	Q8nac5 homo sapien
34	10	1.9	418	2	Q9GRU6_HETSC	Q9gru6 heterodera
35	10	1.9	476	2	O18453_HETGL	O18453 heterodera
36	10	1.9	481	2	Q4PCX0_USTMA	Q4pcx0 ustilago ma
37	10	1.9	570	2	Q9V5S5_DROME	Q9v5s5 drosophila
c 38	10	1.9	578	2	Q8BZY1_MOUSE	Q8bzy1 mus musculu
39	10	1.9	609	2	Q96E25_HUMAN	Q96e25 homo sapien
40	10	1.9	612	2	Q54MS1_DICDI	Q54ms1 dictyosteli
41	10	1.9	642	2	Q7PFH8_ANOGA	Q7pfh8 anopheles g
42	10	1.9	646	2	Q96F45_HUMAN	Q96f45 homo sapien
43	10	1.9	710	1	GSPD2_ERWCH	Q01565 erwinia chr
44	10	1.9	770	2	Q6BIY3_DEBHA	Q6biy3 debaryomyce
45	10	1.9	2063	2	Q4QG78_LEIMA	Q4qg78 leishmania
46	10	1.9	5560	1	SPEN_DROME	Q8sx83 drosophila
47	9	1.7	46	2	Q7UK51_RHOBA	Q7uk51 rhodopirell
48	9	1.7	85	2	Q4YEV7_PLABE	Q4yev7 plasmodium
49	9	1.7	92	2	Q5BQQ1_SCHJA	Q5bqq1 schistosoma
50	9	1.7	109	2	Q71D57_DROYA	Q71d57 drosophila

ALIGNMENTS

RESULT 1
CDSN_HUMAN
ID CDSN_HUMAN STANDARD; PRT; 529 AA.
AC Q15517; O43509; Q5SQ85; Q5STD2; Q7LA70; Q7LA71; Q86Z04; Q8IZU4;
AC Q8IZU5; Q8IZU6; Q8N5P3; Q95IF9; Q9NP52; Q9NPE0; Q9NPG5; Q9NRH4;
AC Q9NRH5; Q9NRH6; Q9NRH7; Q9NRH8; Q9UBH8; Q9UIN6; Q9UIN7; Q9UIN8;
AC Q9UIN9; Q9UIP0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Corneodesmosin precursor (S protein).
GN Name=CDSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LEU-18.
RC TISSUE=Foreskin;
RX MEDLINE=94022396; PubMed=8415725;
RA Zhou Y., Chaplin D.D.;
RT "Identification in the HLA class I region of a gene expressed late in
RT keratinocyte differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9470-9474(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT ALA-408.
RC TISSUE=Epidermis;
RX MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT "Expression cloning of human corneodesmosin proves its identity with
RT the product of the S gene and allows improved characterization of its
RT processing during keratinocyte differentiation.";
RL J. Biol. Chem. 273:22640-22647(1998).
RN [3]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS LEU-18; SER-153 DEL;
RP PHE-202 AND LEU-410.
RX MEDLINE=22256218; PubMed=12366786;
RX DOI=10.1034/j.1399-0039.2002.600110.x;
RA Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT patients with psoriasis.";
RL Tissue Antigens 60:77-83(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RT "Genome diversity in HLA: a new strategy for detection of genetic
RT polymorphisms in expressed genes within the HLA class III and class I
RT regions.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANTS LEU-18;
RP ASN-143; SER-153 DEL; PHE-202; ALA-408; LEU-410 AND ASN-527.
RX MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcott R.,
RA Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLay K.,
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT "The DNA sequence and analysis of human chromosome 6.";
RL Nature 425:805-811(2003).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS LEU-18; SER-153
RP DEL; PHE-202 AND LEU-410.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 30-529 (ALLELES 1.11; 1.21; 1.31;
RP 1.32; 1.41; 1.42; 1.43; 1.51; 1.52; 2.11; 2.21; 2.22 AND 2.23), AND
RP VARIANTS PHE-56; ASN-143 DEL; ASN-150; PHE-202; GLY-401; ALA-408;
RP LEU-410 AND ASN-527.
RX MEDLINE=21108612; PubMed=11169256;
RX DOI=10.1034/j.1399-0039.2001.057001032.x;
RA Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT "Identification of six novel polymorphisms in the human corneodesmosin
RT gene.";
RL Tissue Antigens 57:32-38(2001).
RN [9]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 106-529, VARIANTS ASN-143;
RP PHE-202; GLY-401; ALA-408 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX MEDLINE=20066866; PubMed=10599883;
RX DOI=10.1034/j.1399-0039.1999.540501.x;
RA Jenisch S., Koch S., Henseler T., Nair R.P., Elder J.T., Watts C.E.,
RA Westphal E., Voorhees J.J., Christophers E., Kroenke M.;
RT "Corneodesmosin gene polymorphism demonstrates strong linkage
RT disequilibrium with HLA and association with psoriasis vulgaris.";
RL Tissue Antigens 54:439-449(1999).
RN [10]
RP PARTIAL PROTEIN SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=98058975; PubMed=9395522; DOI=10.1074/jbc.272.50.31770;
RA Simon M., Montezin M., Guerrin M., Durieux J.-J., Serre G.;
RT "Characterization and purification of human corneodesmosin, an
RT epidermal basic glycoprotein associated with corneocyte-specific
RT modified desmosomes.";
RL J. Biol. Chem. 272:31770-31776(1997).
RN [11]
RP INVOLVEMENT IN HTSS.
RX PubMed=12754508; DOI=10.1038/ng1163;
RA Levy-Nissenbaum E., Betz R.C., Frydman M., Simon M., Lahat H.,
RA Bakhan T., Goldman B., Bygum A., Pierick M., Hillmer A.M., Jonca N.,
RA Toribio J., Kruse R., Dewald G., Cichon S., Kubisch C., Guerrin M.,
RA Serre G., Nothen M.M., Pras E.;
RT "Hypotrichosis simplex of the scalp is associated with nonsense
RT mutations in CDSN encoding corneodesmosin.";
RL Nat. Genet. 34:151-153(2003).
RN [12]
RP VARIANTS PHE-202 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX PubMed=10844560; DOI=10.1046/j.1523-1747.2000.00002.x;
RA Enerbaeck C., Enlund F., Inerot A., Samuelsson L., Wahlstroem J.,
RA Swanbeck G., Martinsson T.;
RT "S gene (Corneodesmosin) diversity and its relationship to psoriasis;
RT high content of cSNP in the HLA-linked S gene.";
RL J. Invest. Dermatol. 114:1158-1163(2000).
RN [13]
RP VARIANTS PHE-202; GLY-401; ALA-408; LEU-410 AND ASN-527, AND
RP ASSOCIATION WITH PSORIASIS.
RX PubMed=12472658; DOI=10.1034/j.1399-0039.2002.600403.x;
RA Orru S., Giuressi E., Casula M., Loizedda A., Murru R., Mulargia M.,
RA Masala M.V., Cerimele D., Zucca M., Aste N., Biggio P., Carcassi C.,
RA Contu L.;
RT "Psoriasis is associated with a SNP haplotype of the corneodesmosin
RT gene (CDSN).";
RL Tissue Antigens 60:292-298(2002).
CC -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC intercellular structures that are involved in desquamation.
CC -!- TISSUE SPECIFICITY: Exclusively expressed in skin.
CC -!- POLYMORPHISM: Genetic variation in CDSN may be associated with

```
CC       susceptibility to psoriasis [MIM:177900]. Various CDSN alleles are
CC       known including alleles 1.11, 1.21, 1.31, 1.32, 1.41, 1.42, 1.43,
CC       1.51, 1.52, 2.11, 2.21, 2.22 and 2.23.
CC   -!- DISEASE: Defects in CDSN are a cause of hypotrichosis simplex of
CC       the scalp (HTSS) [MIM:146520]; also known as hypotrichosis Spanish
CC       type. HTSS is an autosomal dominant form of isolated alopecia.
CC       Affected individuals have normal hair in early childhood but
CC       experience progressive loss of scalp hair beginning in the middle
CC       of the first decade and almost complete baldness by the third
CC       decade.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 501.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L20815; AAA21321.1; ALT_SEQ; mRNA.
DR   EMBL; AF030130; AAC24196.1; -; mRNA.

Alignment Scores:
Pred. No.:              1.62e-218       Length:         529
Score:                  245.00          Matches:        345
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            46.4%           Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_HUMAN (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerSerGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 2
CDSN_PANTR
ID   CDSN_PANTR     STANDARD;      PRT;   529 AA.
AC   Q7YR44;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BA000041; BAC78170.1; -; Genomic_DNA.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    529       Corneodesmosin.
FT   COMPBIAS     62    464       Ser-rich.
FT   CARBOHYD    172   172       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;
```

```
Alignment Scores:
Pred. No.:              2.9e-181        Length:         529
Score:                  205.00          Matches:        205
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            38.8%           Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_PANTR (1-529)

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIle 222

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSer 242

Qy        727 TCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 SerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySer 262

Qy        787 GGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysPro 282

Qy        847 TGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySer 302

Qy        907 TCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGC 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 SerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGly 322

Qy        967 TACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGly 342

Qy       1027 CCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCA 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 ProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAla 362

Qy       1087 GCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGly 382

Qy       1147 GGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSer 402

Qy       1207 AGCATTTCCAGCAGC 1221
              ||||||||||||||||
Db        403 SerIleSerSerSer 407

RESULT 3
Q7Z560_HUMAN
ID   Q7Z560_HUMAN PRELIMINARY;      PRT;   133 AA.
AC   Q7Z560;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim Y.-K., Hur S.-S., Kim T.-Y., Kim T.-G., Pyo C.-W.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY328474; AAP94640.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  13187 MW;  350E30DD5AFC652A CRC64;

Alignment Scores:
Pred. No.:              9.25e-79        Length:         133
Score:                  95.00           Matches:        95
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            18.0%           Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q7Z560_HUMAN (1-133)

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIle 58

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSer 78

Qy        727 TCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 SerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySer 98

Qy        787 GGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysPro 118

Qy        847 TGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891
              |||||||||||||||||||||||||||||||||||||||||||||
Db        119 CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGlu 133

RESULT 4
Q7Z559_HUMAN
ID   Q7Z559_HUMAN PRELIMINARY;      PRT;   133 AA.
AC   Q7Z559;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Corneodesmosin (Fragment).
GN   Name=CDSN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Kim Y.-K., Hur S.-S., Kim T.-Y., Kim T.-G., Pyo C.-W.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY328475; AAP94641.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     133    133
SQ   SEQUENCE   133 AA;  13126 MW;  5E6B2CA43FFDB73E CRC64;

Alignment Scores:
Pred. No.:              2.85e-58        Length:         133
Score:                  73.00           Matches:        73
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            13.8%           Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q7Z559_HUMAN (1-133)

Qy        673 TCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCCCAC 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSerHis 80

Qy        733 TCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGTGCC 792
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 SerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGlyAla 100

Qy         793 CCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCC 852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 ProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCysPro 120

Qy         853 CCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAG 891
               |||||||||||||||||||||||||||||||||||||||
Db         121 ProIleThrSerValAspLysSerTyrGlyGlyTyrGlu 133

RESULT 5
CDSN_MACMU
ID   CDSN_MACMU     STANDARD;      PRT;   534 AA.
AC   Q5TM45;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=CDSN;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15269276; DOI=10.1093/molbev/msh216;
RA   Kulski J.K., Anzai T., Shiina T., Inoko H.;
RT   "Rhesus macaque class I duplicon structures, organization, and
RT   evolution within the alpha block of the major histocompatibility
RT   complex.";
RL   Mol. Biol. Evol. 21:2079-2091(2004).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AB128049; BAD69749.1; -; Genomic_DNA.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    534       Corneodesmosin.
FT   COMPBIAS     62    469       Ser-rich.
SQ   SEQUENCE   534 AA;  51983 MW;  CB6A9DF256C32874 CRC64;

Alignment Scores:
Pred. No.:              1.45e-41        Length:         534
Score:                  55.00           Matches:        55
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            10.4%           Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_MACMU (1-534)

Qy         814 CCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGTCCCCCAATCACCTCTGTAGACAAA 873
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         277 ProCysSerAsnGlyGlyLeuProGlyLysProCysProProIleThrSerValAspLys 296

Qy         874 TCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCTGACAGTTATCTGGTTCCAGGCATG 933
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         297 SerTyrGlyGlyTyrGluValValGlyGlySerSerAspSerTyrLeuValProGlyMet 316

Qy         934 ACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTACTTCACCAAA 978
               |||||||||||||||||||||||||||||||||||||||||||||
Db         317 ThrTyrSerLysGlyLysIleTyrProValGlyTyrPheThrLys 331

RESULT 6
CDSN_PIG
ID   CDSN_PIG       STANDARD;      PRT;   225 AA.
AC   O19084;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin (S protein) (Fragment).
GN   Name=CDSN;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   TISSUE=Peripheral blood;
RA   Velten P.W., Rogel-Gaillard C., Renard C., Pontarotti P.,
RA   Tazi-Ahnini R., Vaiman M., Chardon P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; Z97402; CAB10690.1; -; Genomic_DNA.
FT   COMPBIAS     79     85       Poly-Ser.
FT   COMPBIAS    129    140       Poly-Ser.
FT   COMPBIAS    159    168       Poly-Ser.
FT   COMPBIAS    179    184       Poly-Ser.
FT   NON_TER       1      1
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  21644 MW;  8544070C41F7B280 CRC64;

Alignment Scores:
Pred. No.:              1e-11           Length:         225
Score:                  23.00           Matches:        23
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            4.4%            Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_PIG (1-225)

Qy        1009 CCTTCCTTTGCAGCTGGGCCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATC 1068
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          57 ProSerPheAlaAlaGlyProProIleSerGluGlyLysTyrPheSerSerAsnProIle 76

Qy        1069 ATCCCCAGC 1077
               |||||||||
Db          77 IleProSer 79

RESULT 7
CDSN_MOUSE
ID   CDSN_MOUSE     STANDARD;      PRT;   561 AA.
AC   Q7TPC1;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=Cdsn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C3H/He; TISSUE=Osteoblast;
```

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BC055373; AAH55373.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039518; Mus musculus.
DR   MGI; MGI:3505689; Cdsn.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    561       Corneodesmosin.
FT   COMPBIAS     62    493       Ser-rich.
SQ   SEQUENCE   561 AA;  54414 MW;  7A6EBEF364038746 CRC64;

Alignment Scores:
Pred. No.:              4.15e-07        Length:         561
Score:                  18.00           Matches:        18
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            3.4%            Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x CDSN_MOUSE (1-561)

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 SerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 56

RESULT 8
Q767M8_PIG
ID   Q767M8_PIG  PRELIMINARY;      PRT;   262 AA.
AC   Q767M8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin.
GN   Name=CDSN;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=large white;
RA   Shigenari A., Ando A., Renard C., Chardon P., Shiina T., Kulski J.K.,
RA   Yasue H., Inoko H.;
RT   "Nucleotide sequencing analysis of the swine 433-kb genomic segment
RT   located between the non-classical and classical SLA class I gene
RT   clusters.";
RL   Immunogenetics 55:695-705(2004).
DR   EMBL; AB113354; BAD08421.1; -; Genomic_DNA.
SQ   SEQUENCE   262 AA;  25289 MW;  21F44CFE54C920F2 CRC64;

Alignment Scores:
Pred. No.:              0.173           Length:         262
Score:                  12.00           Matches:        12
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.3%            Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q767M8_PIG (1-262)

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCG 702
              |||||||||||||||||||||||||||||||||||
Db        197 ProAspSerProCysSerGlyGlyProIleValSer 208

RESULT 9
GAP_CAEEL
ID   GAP_CAEEL      STANDARD;      PRT;  1317 AA.
AC   P34288; P34588; P34589; Q8ST20; Q9TY64;
DT   01-FEB-1994 (Rel. 28, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   GTPase-activating protein GAP (CeGAP).
GN   Name=gap; ORFNames=C04D8.1, ZC21.4/ZC21.5;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND FUNCTION.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94117490; PubMed=8288633;
RA   Chen W., Blanc J., Lim L.;
RT   "Characterization of a promiscuous GTPase-activating protein that has
RT   a bcr-related domain from Caenorhabditis elegans.";
RL   J. Biol. Chem. 269:820-823(1994).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398; DOI=10.1038/368032a0;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
RN   [4]
RP   SEQUENCE REVISION.
RG   WormBase consortium;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: GTPase-activating protein for members of the Rho
```

```
CC       subfamily including Rac1, RhoA and cdc42 and other Ras-related
CC       subfamiles including let-60.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Contains 1 PH domain.
CC   -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U02289; AAA18934.1; ALT_INIT; mRNA.
DR   EMBL; L16687; AAK71357.2; -; Genomic_DNA.
DR   PIR; S44746; S44746.
DR   PIR; S44877; S44877.
DR   HSSP; Q07960; 1RGP.
DR   Ensembl; C04D8.1; Caenorhabditis elegans.
DR   WormBase; WBGene00015418; C04D8.1.
DR   WormPep; C04D8.1; CE33028.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005099; F:Ras GTPase activator activity; IDA.
DR   GO; GO:0005100; F:Rho GTPase activator activity; IDA.
DR   GO; GO:0007265; P:Ras protein signal transduction; IDA.
DR   GO; GO:0007266; P:Rho protein signal transduction; IDA.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR011993; PH_type.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   Complete proteome; GTPase activation.
FT   DOMAIN      311    438       PH.
FT   DOMAIN      660    858       Rho-GAP.
FT   COMPBIAS    106    264       Ser-rich.
FT   COMPBIAS    496    513       Gln-rich.
FT   COMPBIAS   1064   1088       Ala-rich.
FT   CONFLICT    276    276       L -> I (in Ref. 1).
SQ   SEQUENCE   1317 AA;  142071 MW;  5871763017001A39 CRC64;

Alignment Scores:
Pred. No.:              0.148          Length:        1317
Score:                  12.00          Matches:       12
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            2.3%           Indels:        0
DB:                     1              Gaps:          0

US-10-677-865A-2 (1-1590) x GAP_CAEEL (1-1317)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              |||||||||||||||||||||||||||||||||||||
Db        634 SerGlyGlyGlySerSerGlySerSerSerGlySer 645

RESULT 10
Q60YS0_CAEBR
ID   Q60YS0_CAEBR PRELIMINARY;      PRT;  1344 AA.
AC   Q60YS0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein CBG18119 (Fragment).
GN   Name=CBG18119;
OS   Caenorhabditis briggsae.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RG   The C.briggsae Sequencing Consortium;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 PH domain.
DR   EMBL; CAAC01000088; CAE71244.1; -; Genomic_DNA.
DR   InterPro; IPR001849; PH.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00169; PH; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00233; PH; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   1344 AA;  144974 MW;  C1043153AECB059C CRC64;

Alignment Scores:
Pred. No.:              0.148          Length:        1344
Score:                  12.00          Matches:       12
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            2.3%           Indels:        0
DB:                     2              Gaps:          0

US-10-677-865A-2 (1-1590) x Q60YS0_CAEBR (1-1344)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              |||||||||||||||||||||||||||||||||||||
Db        654 SerGlyGlyGlySerSerGlySerSerSerGlySer 665

RESULT 11
Q7PFN9_ANOGA
ID   Q7PFN9_ANOGA PRELIMINARY;      PRT;   208 AA.
AC   Q7PFN9;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ENSANGP00000024119 (Fragment).
GN   ORFNames=ENSANGG00000020169;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC   Anophelinae; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PEST;
RG   The Anopheles gambiae Sequence Committee;
RT   "Anopheles gambiae re-annotation.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PEST;
RG   The Anopheles gambiae Sequence Committee;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008839; EAA45240.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  21743 MW;  419504FF02B89D2F CRC64;

Alignment Scores:
Pred. No.:              1.51           Length:        208
Score:                  11.00          Matches:       11
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            2.1%           Indels:        0
DB:                     2              Gaps:          0

US-10-677-865A-2 (1-1590) x Q7PFN9_ANOGA (1-208)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
```

```
                    ||||||||||||||||||||||||||||||||
Db        132 GlyGlySerSerGlySerSerSerGlySerSer 142

RESULT 12
Q8MYJ1_DICDI
ID   Q8MYJ1_DICDI PRELIMINARY;      PRT;   262 AA.
AC   Q8MYJ1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Similar to Haliotis rufescens (California red abalone). LUSTRIN A
DE   (Hypothetical protein).
GN   ORFNames=DDB0169451;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RX   MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA   Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA   Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA   Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL   Nature 418:79-85(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Baumgart C.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
DR   EMBL; AC116979; AAM34319.3; -; Genomic_DNA.
DR   EMBL; AAFI01000032; EAL69324.1; -; Genomic_DNA.
DR   InterPro; IPR007117; Expan_Lol_pI_C.
KW   Hypothetical protein.
SQ   SEQUENCE   262 AA;  24798 MW;  F8241CAA153D622B CRC64;

Alignment Scores:
Pred. No.:              1.48          Length:         262
Score:                  11.00         Matches:        11
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            2.1%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q8MYJ1_DICDI (1-262)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
                ||||||||||||||||||||||||||||||||
Db        135 GlyGlySerSerGlySerSerSerGlySerSer 145

RESULT 13
Q7XJL9_ARATH
ID   Q7XJL9_ARATH PRELIMINARY;      PRT;   698 AA.
AC   Q7XJL9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   At2g17010 protein.
GN   Name=At2g17010;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC002354; AAF18596.1; -; Genomic_DNA.
DR   PIR; A84547; A84547.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR006685; MscS_channel.
DR   Pfam; PF00924; MS_channel; 1.
SQ   SEQUENCE   698 AA;  80007 MW;  F94A39EC8C78DF84 CRC64;

Alignment Scores:
Pred. No.:              1.35          Length:         698
Score:                  11.00         Matches:        11
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            2.1%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q7XJL9_ARATH (1-698)

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
                ||||||||||||||||||||||||||||||||
Db        196 SerSerSerSerHisSerSerSerSerSerSer 206

RESULT 14
Q8SX98_DROME
ID   Q8SX98_DROME PRELIMINARY;      PRT;   708 AA.
AC   Q8SX98;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   GH25793p (CG13503-PF, isoform F).
GN   ORFNames=CG13503;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
```

RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY094731; AAM11084.1; -; mRNA.
DR EMBL; AE003456; AAM68204.1; -; Genomic_DNA.
DR Ensembl; CG13503; Drosophila melanogaster.
DR FlyBase; FBgn0034695; CG13503.
DR GO; GO:0007015; P:actin filament organization; IMP.
DR GO; GO:0008360; P:regulation of cell shape; IMP.
DR InterPro; IPR003124; WH2.
DR Pfam; PF02205; WH2; 1.
DR SMART; SM00246; WH2; 1.
SQ SEQUENCE 708 AA; 71215 MW; 82457AD10E609F8B CRC64;

Alignment Scores:

Pred. No.:	1.34	Length:	708
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x Q8SX98_DROME (1-708)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||
Db        325 GlyGlySerSerGlySerSerSerGlySerSer 335
```

RESULT 15
TMCC2_HUMAN
ID TMCC2_HUMAN STANDARD; PRT; 709 AA.
AC O75069; Q6ZN09;
DT 15-DEC-1998 (Rel. 37, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane and coiled-coil domains protein 2 (Cerebral protein 11).
GN Name=TMCC2; Synonyms=KIAA0481; ORFNames=hucep-11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,

```
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 60-709.
RC   TISSUE=Brain;
RX   MEDLINE=98116662; PubMed=9455484;
RA   Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
RA   Nakajima D., Nomura N., Ohara O.;
RT   "Characterization of cDNA clones in size-fractionated cDNA libraries
RT   from human brain.";
RL   DNA Res. 4:345-349(1997).
RN   [3]
RP   SEQUENCE REVISION.
RX   MEDLINE=22158633; PubMed=12168954;
RA   Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT   "Construction of expression-ready cDNA clones for KIAA genes: manual
RT   curation of 330 KIAA cDNA clones.";
RL   DNA Res. 9:99-106(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [MRNA] OF 226-709.
RC   TISSUE=Brain;
RA   Yoshimoto M., Yazaki M., Takayama K., Matsumoto K., Tsuritani K.;
RT   "Molecular cloning of a gene specifically expressed in human central
RT   nervous system.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the TEX28 family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AK131419; BAD18566.1; -; mRNA.
DR   EMBL; AB007950; BAA32326.2; -; mRNA.
DR   EMBL; AB001596; BAB46924.1; -; mRNA.
DR   Ensembl; ENSG00000133069; Homo sapiens.
DR   HGNC; HGNC:24239; TMCC2.
KW   Coiled coil; Transmembrane.
FT   TRANSMEM    649    669       Potential.
FT   TRANSMEM    682    702       Potential.
FT   COILED      333    368       Potential.
FT   COILED      514    633       Potential.
FT   CONFLICT     60     69       GPRSKPPDLK -> AAQSQEVRAQ (in Ref. 2).
FT   CONFLICT    226    249       TALLLADGSNVYLLAEEAEGIGDK -> MNQVVQPLMSRHS
FT                                ACRGSQAHLSW (in Ref. 4).
SQ   SEQUENCE   709 AA;  77449 MW;  002993EA6B0E4546 CRC64;

Alignment Scores:
Pred. No.:                1.34          Length:          709
Score:                    11.00         Matches:         11
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              2.1%          Indels:          0
DB:                       1             Gaps:            0

US-10-677-865A-2 (1-1590) x TMCC2_HUMAN (1-709)

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||||||
Db        170 SerSerGlyGlyGlySerSerGlySerSerSer 180

RESULT 16
Q8MLU0_DROME
ID   Q8MLU0_DROME PRELIMINARY;      PRT;   751 AA.
AC   Q8MLU0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG13503-PA, isoform A (Cg13503-pb, isoform b).
GN   Name=CG13503; ORFNames=CG13503;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
```

```
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- INTERACTION:
CC       Q9VPU1:dock; NbExp=1; IntAct=EBI-157685, EBI-189398;
CC       Q9VAT0:WASp; NbExp=1; IntAct=EBI-157685, EBI-132019;
DR   EMBL; AE003456; AAM68205.1; -; Genomic_DNA.
DR   IntAct; Q8MLU0; -.
DR   Ensembl; CG13503; Drosophila melanogaster.
DR   FlyBase; FBgn0034695; CG13503.
DR   GO; GO:0007015; P:actin filament organization; IMP.
DR   GO; GO:0008360; P:regulation of cell shape; IMP.
DR   InterPro; IPR003124; WH2.
DR   Pfam; PF02205; WH2; 1.
DR   SMART; SM00246; WH2; 1.
SQ   SEQUENCE   751 AA;  76385 MW;  9F4345952EC7CA9D CRC64;

Alignment Scores:
Pred. No.:              1.34          Length:         751
Score:                  11.00         Matches:        11
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            2.1%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q8MLU0_DROME (1-751)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||
Db        325 GlyGlySerSerGlySerSerSerGlySerSer 335

RESULT 17
Q7KVL6_DROME
ID   Q7KVL6_DROME PRELIMINARY;      PRT;   761 AA.
AC   Q7KVL6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG13503-PE, isoform E.
GN   Name=CG13503; ORFNames=CG13503;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
```

```
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003456; AAM68208.2; -; Genomic_DNA.
DR   Ensembl; CG13503; Drosophila melanogaster.
DR   InterPro; IPR003124; WH2.
DR   Pfam; PF02205; WH2; 1.
DR   SMART; SM00246; WH2; 1.
SQ   SEQUENCE   761 AA;  77628 MW;  1AF39E27997F3B5A CRC64;

Alignment Scores:
Pred. No.:              1.33          Length:        761
Score:                  11.00         Matches:       11
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            2.1%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q7KVL6_DROME (1-761)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||
Db        335 GlyGlySerSerGlySerSerSerGlySerSer 345

RESULT 18
Q5K4K9_CHLRE
ID   Q5K4K9_CHLRE PRELIMINARY;      PRT;  2103 AA.
AC   Q5K4K9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative group II intron trans-splicing factor Raa1 precursor.
GN   Name=raa1;
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC   Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Goldschmidt-Clermont M., Perron K., Rochaix J.D.;
RT   "Raa1, a novel factor involved in trans-splicing of a group II intron
RT   in the chloroplast of Chlamydomonas reinhardtii.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ605114; CAE53330.1; -; mRNA.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   InterPro; IPR000637; A+T_hook.
DR   InterPro; IPR002086; Aldehyd_dehydrog.
DR   Pfam; PF02178; AT_hook; 2.
DR   PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW   Transit peptide.
FT   TRANSIT       1     54       Potential.
FT   CHAIN        55   2103       putative group II intron trans-splicing
FT                                 factor Raa1.
SQ   SEQUENCE   2103 AA;  206221 MW;  44872D2CB011FCA2 CRC64;

Alignment Scores:
Pred. No.:              1.21          Length:        2103
Score:                  11.00         Matches:       11
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            2.1%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5K4K9_CHLRE (1-2103)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGCATT 291
              ||||||||||||||||||||||||||||||||
Db        781 GlySerSerGlySerSerSerGlySerSerIle 791

RESULT 19
Q6MAJ1_PARUW
ID   Q6MAJ1_PARUW PRELIMINARY;      PRT;    98 AA.
AC   Q6MAJ1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=pc1684;
OS   Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
OC   Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX   NCBI_TaxID=264201;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15073324;
RA   Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA   Fartmann B., Brandt P., Nyakatura G.J., Droege M., Frishman D.,
RA   Rattei T., Mewes H.-W., Wagner M.;
RT   "Illuminating the evolutionary history of chlamydiae.";
RL   Science 304:728-730(2004).
DR   EMBL; BX908798; CAF24408.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   98 AA;  10303 MW;  DA4FF65F6A18D406 CRC64;

Alignment Scores:
Pred. No.:              13.9          Length:        98
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q6MAJ1_PARUW (1-98)

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||||
Db         78 SerSerSerSerHisSerSerSerSerSer 87

RESULT 20
Q6MG23_RAT
ID   Q6MG23_RAT  PRELIMINARY;      PRT;   129 AA.
AC   Q6MG23;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Corneodesmosin.
GN   Name=Cdsn;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Brown Norway;
RX   PubMed=15060004; DOI=10.1101/gr.1987704;
RA   Hurt P., Walter L., Sudbrak R., Klages S., Mueller I., Shiina T.,
RA   Inoko H., Lehrach H., Guenther E., Reinhardt R., Himmelbauer H.;
RT   "The genomic sequence and comparative analysis of the rat major
RT   histocompatibility complex.";
RL   Genome Res. 14:631-639(2004).
```

```
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Brown Norway;
RA   Boehm S., Borzym K., Gelling S., Gimmel V., Heitmann K., Kosiura A.,
RA   Lang N., Lehrack S., Thiel J., Sontag M., Hurt P., Himmelbauer H.,
RA   Sudbrak R., Reinhardt R.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Brown Norway;
RA   Shiina T., Inoko H.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX883047; CAE84024.1; -; Genomic_DNA.
DR   InterPro; IPR001917; Aminotrans_II.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
SQ   SEQUENCE   129 AA;  13729 MW;  13B7791C0A25AFF2 CRC64;

Alignment Scores:
Pred. No.:              13.5          Length:         129
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q6MG23_RAT (1-129)

Qy         79 CTGCCAGGGACCTTGGCTAAGAGCATTGGC 108
              ||||||||||||||||||||||||||||||
Db         27 LeuProGlyThrLeuAlaLysSerIleGly 36

RESULT 21
Q4RZX2_TETNG
ID   Q4RZX2_TETNG PRELIMINARY;      PRT;   158 AA.
AC   Q4RZX2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 18 SCAF14786, whole genome shotgun sequence.
GN   ORFNames=GSTENG00026292001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014786; CAG06060.1; -; Genomic_DNA.
SQ   SEQUENCE   158 AA;  17863 MW;  AAD531103AF462F9 CRC64;

Alignment Scores:
Pred. No.:              13.2          Length:         158
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q4RZX2_TETNG (1-158)

Qy         259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
               ||||||||||||||||||||||||||||||
Db         141 GlySerSerGlySerSerSerGlySerSer 150

RESULT 22
Q569P6_BRARE
ID   Q569P6_BRARE PRELIMINARY;      PRT;   173 AA.
AC   Q569P6;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein zgc:110549.
GN   Name=zgc:110549;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RG   NIH MGC Project;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC092359; AAH92359.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   173 AA;  19646 MW;  69F2299BD929E5A6 CRC64;

Alignment Scores:
Pred. No.:              13.1          Length:         173
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q569P6_BRARE (1-173)

Qy         259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
               ||||||||||||||||||||||||||||||
```

```
Db        157 GlySerSerGlySerSerSerGlySerSer 166

RESULT 23
Q21636_CAEEL
ID   Q21636_CAEEL PRELIMINARY;      PRT;   206 AA.
AC   Q21636;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein R02D5.1.
GN   Name=R02D5.1; ORFNames=R02D5.1;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Bristol N2;
RA   Matthews L.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
CC   -!- INTERACTION:
CC       Q9U9Y8:lit-1; NbExp=1; IntAct=EBI-329142, EBI-318513;
DR   EMBL; Z78015; CAB01433.1; -; Genomic_DNA.
DR   PIR; T23852; T23852.
DR   IntAct; Q21636; -.
DR   Ensembl; R02D5.1; Caenorhabditis elegans.
DR   WormBase; WBGene00010976; R02D5.1.
DR   WormPep; R02D5.1; CE12486.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   206 AA;  22317 MW;  F6E5448ADBA26F48 CRC64;

Alignment Scores:
Pred. No.:              12.9          Length:        206
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q21636_CAEEL (1-206)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db         84 SerSerSerHisSerSerSerSerSerSer 93

RESULT 24
Q9R3S1_ACTPL
ID   Q9R3S1_ACTPL PRELIMINARY;      PRT;   364 AA.
AC   Q9R3S1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Outer membrane lipoprotein precursor.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=femoe, and S1421;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Femo, and S1421;
RA   Ito H., Osaki M., Uchida I., Ohya T., Sekizaki T.;
RT   "Comparative study of nucleotide sequences of genes for outer membrane
RT   lipoprotein of Actinobacillus pleuropneumoniae.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86681; AAD00614.1; -; Genomic_DNA.
DR   EMBL; AB007578; BAD99427.1; -; Genomic_DNA.
DR   EMBL; U86677; AAD00610.1; -; Genomic_DNA.
DR   EMBL; AB007574; BAD99423.1; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein; Signal.
FT   SIGNAL          1     19       Potential.
SQ   SEQUENCE   364 AA;  39576 MW;  4A82E539652485A9 CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:        364
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9R3S1_ACTPL (1-364)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 25
OMLA_ACTPL
ID   OMLA_ACTPL     STANDARD;      PRT;   365 AA.
AC   Q02937;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Outer membrane lipoprotein A precursor.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   STRAIN=Serotype 1 / Isolate AP37;
RX   MEDLINE=93138779; PubMed=8423086;
RA   Gerlach G.-F., Anderson C., Klashinsky S., Rossi-Campos A.,
RA   Potter A.A., Willson P.J.;
RT   "Molecular characterization of a protective outer membrane lipoprotein
RT   (OmlA) from Actinobacillus pleuropneumoniae serotype 1.";
RL   Infect. Immun. 61:565-572(1993).
CC   -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC       anchor.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; L06318; AAC41456.1; -; Genomic_DNA.
DR   PIR; A49235; A49235.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Lipoprotein; Membrane; Outer membrane; Palmitate; Signal.
FT   SIGNAL        1     19
FT   CHAIN        20    365       Outer membrane lipoprotein A.
FT   LIPID        20     20       N-palmitoyl cysteine (Probable).
FT   LIPID        20     20       S-diacylglycerol cysteine (Probable).
```

```
SQ   SEQUENCE   365 AA;  39784 MW;  20D4B3FAEBE2842F CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:        365
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     1             Gaps:          0

US-10-677-865A-2 (1-1590) x OMLA_ACTPL (1-365)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 26
Q9R3L8_ACTPL
ID   Q9R3L8_ACTPL PRELIMINARY;      PRT;   365 AA.
AC   Q9R3L8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Outer membrane lipoprotein precursor.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=56153, S4074, and CVI 13261;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CVJ13261, 56153, and 4074;
RA   Ito H., Osaki M., Uchida I., Ohya T., Sekizaki T.;
RT   "Comparative study of nucleotide sequences of genes for outer membrane
RT   lipoprotein of Actinobacillus pleuropneumoniae.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86686; AAD00619.1; -; Genomic_DNA.
DR   EMBL; AB007581; BAD99429.1; -; Genomic_DNA.
DR   EMBL; AB007583; BAD99431.1; -; Genomic_DNA.
DR   EMBL; U86675; AAD00608.1; -; Genomic_DNA.
DR   EMBL; U86684; AAD00617.1; -; Genomic_DNA.
DR   EMBL; AB007572; BAD99421.1; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein; Signal.
FT   SIGNAL        1     19       Potential.
SQ   SEQUENCE   365 AA;  39842 MW;  5E46DFCBB542781E CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:        365
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9R3L8_ACTPL (1-365)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 27
Q9RCG1_ACTPL
ID   Q9RCG1_ACTPL PRELIMINARY;      PRT;   365 AA.
AC   Q9RCG1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Outer membrane lipoprotein.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=8329;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=8329;
RA   Ito H., Osaki M., Uchida I., Ohya T., Sekizaki T.;
RT   "Comparative study of nucleotide sequences of genes for outer membrane
RT   lipoprotein of Actinobacillus pleuropneumoniae.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86687; AAD00620.1; -; Genomic_DNA.
DR   EMBL; AB007584; BAD99432.1; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein.
SQ   SEQUENCE   365 AA;  39750 MW;  E20A60DBB246DAE5 CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:        365
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9RCG1_ACTPL (1-365)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 28
Q9RIH8_ACTPL
ID   Q9RIH8_ACTPL PRELIMINARY;      PRT;   368 AA.
AC   Q9RIH8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Outer membrane lipoprotein.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=405;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=405;
RA   Gram T., Ahrens P., Angen O.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86683; AAD00616.2; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
```

```
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein.
SQ   SEQUENCE   368 AA;  40031 MW;  9584C0ED96E6CC9A CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:         368
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q9RIH8_ACTPL (1-368)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 29
Q9RCG3_ACTPL
ID   Q9RCG3_ACTPL PRELIMINARY;      PRT;   368 AA.
AC   Q9RCG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Outer membrane lipoprotein precursor.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=S1536;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=S1536, and 405;
RA   Ito H., Osaki M., Uchida I., Ohya T., Sekizaki T.;
RT   "Comparative study of nucleotide sequences of genes for outer membrane
RT   lipoprotein of Actinobacillus pleuropneumoniae.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86676; AAD00609.1; -; Genomic_DNA.
DR   EMBL; AB007573; BAD99422.1; -; Genomic_DNA.
DR   EMBL; AB007580; BAD99428.1; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein; Signal.
FT   SIGNAL        1     19       Potential.
SQ   SEQUENCE   368 AA;  39967 MW;  0281948BF2E8E783 CRC64;

Alignment Scores:
Pred. No.:              12.2          Length:         368
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q9RCG3_ACTPL (1-368)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 30
YB1E_SCHPO
ID   YB1E_SCHPO     STANDARD;      PRT;   374 AA.
AC   P87179; Q9USB7; Q9USD8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Serine-rich protein C30B4.01c precursor.
GN   ORFNames=SPBC30B4.01c, SPBC3D6.14c;
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA   Sgouros J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J.R., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Duesterhoeft A., Fritzc C., Holzer E., Moestl D.,
RA   Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA   Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 31-66 AND 136-345,
RP   AND SUBCELLULAR LOCATION.
RC   STRAIN=968 h90;
RX   MEDLINE=20223868; PubMed=10759889;
RA   Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA   Hiraoka Y.;
RT   "Large-scale screening of intracellular protein localization in living
RT   fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL   Genes Cells 5:169-190(2000).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated.
CC   -!- SIMILARITY: Contains 1 WSC domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AL031262; CAA20314.2; -; Genomic_DNA.
DR   EMBL; Z95620; CAB09116.1; -; Genomic_DNA.
DR   EMBL; AB027834; BAA87138.1; -; Genomic_DNA.
DR   EMBL; AB027890; BAA87194.1; -; Genomic_DNA.
DR   PIR; T40167; T40167.
DR   GeneDB_Spombe; SPBC30B4.01c; -.
DR   InterPro; IPR002889; WSC_carb_bd.
DR   Pfam; PF01822; WSC; 1.
DR   SMART; SM00321; WSC; 1.
KW   Complete proteome; Glycoprotein; Signal; Transmembrane.
FT   SIGNAL        1     29       Potential.
FT   CHAIN        30    374       Serine-rich protein C30B4.01c.
FT   TRANSMEM    293    313       Potential.
FT   DOMAIN       31    119       WSC.
```

```
FT   COMPBIAS    128    271       Ser-rich.
FT   CARBOHYD    278    278       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    284    284       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   374 AA;  38398 MW;  C7942A0E0E47941E CRC64;

Alignment Scores:
Pred. No.:                12.2          Length:        374
Score:                    10.00         Matches:       10
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.9%          Indels:        0
DB:                       1             Gaps:          0

US-10-677-865A-2 (1-1590) x YB1E_SCHPO (1-374)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        215 SerSerSerHisSerSerSerSerSerSer 224

RESULT 31
Q9RCG2_ACTPL
ID   Q9RCG2_ACTPL PRELIMINARY;      PRT;   375 AA.
AC   Q9RCG2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Outer membrane lipoprotein precursor.
GN   Name=omlA;
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=M62;
RX   MEDLINE=98126231; PubMed=9466755;
RA   Gram T., Ahrens P.;
RT   "Improved diagnostic PCR assay for Actinobacillus pleuropneumoniae
RT   based on the nucleotide sequence of an outer membrane lipoprotein.";
RL   J. Clin. Microbiol. 36:443-448(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=M62;
RA   Ito H., Osaki M., Uchida I., Ohya T., Sekizaki T.;
RT   "Comparative study of nucleotide sequences of genes for outer membrane
RT   lipoprotein of Actinobacillus pleuropneumoniae.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U86678; AAD00611.1; -; Genomic_DNA.
DR   EMBL; AB007575; BAD99424.1; -; Genomic_DNA.
DR   InterPro; IPR005014; Lipoprotein_14.
DR   Pfam; PF03346; Lipoprotein_14; 1.
KW   Lipoprotein; Signal.
FT   SIGNAL      1     19       Potential.
SQ   SEQUENCE   375 AA;  40840 MW;  67BA62BE58E01A02 CRC64;

Alignment Scores:
Pred. No.:                12.2          Length:        375
Score:                    10.00         Matches:       10
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.9%          Indels:        0
DB:                       2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9RCG2_ACTPL (1-375)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 32
Q9A8B9_CAUCR
ID   Q9A8B9_CAUCR PRELIMINARY;      PRT;   375 AA.
AC   Q9A8B9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein CC1435.
GN   OrderedLocusNames=CC1435;
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA   Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA   Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA   Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005818; AAK23416.1; -; Genomic_DNA.
DR   PIR; D87427; D87427.
DR   TIGR; CC1435; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   InterPro; IPR002575; APH_trans.
DR   InterPro; IPR008266; Tyr_pkinase_AS.
DR   Pfam; PF01636; APH; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   375 AA;  40552 MW;  2DBD2A404F878CD4 CRC64;

Alignment Scores:
Pred. No.:                12.2          Length:        375
Score:                    10.00         Matches:       10
Percent Similarity:       100.0%        Conservative:  0
Best Local Similarity:    100.0%        Mismatches:    0
Query Match:              1.9%          Indels:        0
DB:                       2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9A8B9_CAUCR (1-375)

Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCT 481
              ||||||||||||||||||||||||||||||
Db        121 AlaAlaLeuIleArgAlaAlaAlaAlaAla 130

RESULT 33
Q8NAC5_HUMAN
ID   Q8NAC5_HUMAN PRELIMINARY;      PRT;   406 AA.
AC   Q8NAC5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein FLJ35578.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
```

```
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
DR   EMBL; AK092897; BAC03999.1; -; mRNA.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ   SEQUENCE   406 AA;  39681 MW;  68679FBFA848EB87 CRC64;

Alignment Scores:
Pred. No.:              12.1          Length:         406
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q8NAC5_HUMAN (1-406)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81

RESULT 34
Q9GRU6_HETSC
ID   Q9GRU6_HETSC PRELIMINARY;      PRT;   418 AA.
AC   Q9GRU6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Beta-1,4-endoglucanase 1 (Fragment).
GN   Name=eng1;
OS   Heterodera schachtii.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC   Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
OX   NCBI_TaxID=97005;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   De Meutter J., Tytgat T., Gheysen G.D.C., Coomans A., Gheysen G.D.R.;
RT   "Biochemical characterisation of pharyngeal secretions from Heterodera
RT   schachtii stage 2 juveniles.";
RL   Eur. J. Biochem. 0:0-0(0).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   De Meutter J.J.J.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ299386; CAC12958.1; -; mRNA.
DR   HSSP; P07103; 1EGZ.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0030247; F:polysaccharide binding; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR001919; Bac_cellose-bind.
DR   InterPro; IPR012291; Carb_Cell_bind.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF00150; Cellulase; 1.
DR   SMART; SM00637; CBD_II; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   418 AA;  43739 MW;  EAA16B402EDC5A9C CRC64;

Alignment Scores:
Pred. No.:              12.1          Length:         418
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q9GRU6_HETSC (1-418)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        294 GlySerSerGlySerSerSerGlySerSer 303

RESULT 35
O18453_HETGL
ID   O18453_HETGL PRELIMINARY;      PRT;   476 AA.
AC   O18453;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Beta-1,4-endoglucanase-1 precursor (EC 3.2.1.4).
GN   Name=HG-eng-1;
OS   Heterodera glycines (Soybean cyst nematode).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC   Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
OX   NCBI_TaxID=51029;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98226740; PubMed=9560201; DOI=10.1073/pnas.95.9.4906;
RA   Smant G., Stokkermans J.P.W.G., Yan Y., De Boer J.M., Baum T.J.,
RA   Wang X., Hussey R.S., Gommers F.J., Henrissat B., Davis E.L.,
RA   Helder J., Schots A., Bakker J.;
RT   "Endogenous cellulases in animals: isolation of beta-1, 4-
RT   endoglucanase genes from two species of plant-parasitic cyst
RT   nematodes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:4906-4911(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99361199;
RA   de Boer J.M., Yan Y., Wang X., Smant G., Hussey R.S., Davis E.L.,
RA   Baum T.J.;
RT   "Developmental expression of secretory beta-1,4-endoglucanases in the
RT   subventral esophageal glands of Heterodera glycines.";
RL   Mol. Plant Microbe Interact. 12:663-669(1999).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98440369; PubMed=9767113; DOI=10.1016/S0378-1119(98)00413-2;
RA   Yan Y., Smant G., Stokkermans J., Qin L., Helder J., Baum T.,
RA   Schots A., Davis E.;
RT   "Genomic organization of four beta-1,4-endoglucanase genes in plant-
RT   parasitic cyst nematodes and its evolutionary implications.";
RL   Gene 220:61-70(1998).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=OP50;
RA   Yan Y., Smant G., Stokkermans J.W.P.G., Wang X., Hussey R.S.,
RA   Bakker J., Helder J., Schots A., Davis E.L.;
RT   "Genomic sequence of soybean cyst nematode cellulase HG-eng-1.";
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF006052; AAC48327.1; -; mRNA.
DR   EMBL; AF052733; AAC15707.1; -; Genomic_DNA.
```

```
DR   HSSP; P07103; 1EGZ.
DR   GO; GO:0008810; F:cellulase activity; IEA.
DR   GO; GO:0030247; F:polysaccharide binding; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR001919; Bac_cellose-bind.
DR   InterPro; IPR012291; Carb_Cell_bind.
DR   InterPro; IPR001547; Glyco_hydro_5.
DR   Pfam; PF00150; Cellulase; 1.
DR   SMART; SM00637; CBD_II; 1.
KW   Glycosidase; Hydrolase; Signal.
FT   SIGNAL       1     27       Potential.
FT   CHAIN       28    476       beta-1,4-endoglucanase-1.
SQ   SEQUENCE   476 AA;  49815 MW;  2056F90CA0783FF7 CRC64;

Alignment Scores:
Pred. No.:              11.9          Length:         476
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x O18453_HETGL (1-476)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        344 GlySerSerGlySerSerSerGlySerSer 353

RESULT 36
Q4PCX0_USTMA
ID   Q4PCX0_USTMA PRELIMINARY;      PRT;   481 AA.
AC   Q4PCX0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=UM02043.1;
OS   Ustilago maydis 521.
OC   Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC   Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX   NCBI_TaxID=237631;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=521;
RA   Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA   Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA   Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA   Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA   Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA   Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA   Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA   David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA   Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA   Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA   Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA   Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA   Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA   Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA   Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA   Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA   Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA   Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA   Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA   Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA   Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA   Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA   Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA   Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA   O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA   Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA   Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA   Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA   Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA   Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA   Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA   Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA   Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA   Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA   Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA   Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA   Zimmer A., Zody M., Lander E.;
RT   "The genome sequence of Ustilago maydis.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACP01000074; EAK83095.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   481 AA;  49080 MW;  29EC48BA8C08FB0A CRC64;

Alignment Scores:
Pred. No.:              11.9          Length:         481
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q4PCX0_USTMA (1-481)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        443 GlySerSerGlySerSerSerGlySerSer 452

RESULT 37
Q9V5S5_DROME
ID   Q9V5S5_DROME PRELIMINARY;      PRT;   570 AA.
AC   Q9V5S5; Q8MR37;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   CG33473-PB (Transcription factor DKLF) (HL07808p).
GN   Name=luna; Synonyms=CG17326; ORFNames=CG33473;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```

```
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22890060; PubMed=14527717; DOI=10.1016/S0378-1119(03)00720-0;
RA   De Graeve F., Smaldone S., Laub F., Mlodzik M., Bhat M., Ramirez F.;
RT   "Identification of the Drosophila progenitor of mammalian Krueppel-
RT   like factors 6 and 7 and a determinant of fly development.";
RL   Gene 314:55-62(2003).
RN   [8]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003827; AAF58723.3; -; Genomic_DNA.
DR   EMBL; AF461497; AAO85420.1; -; mRNA.
DR   EMBL; AY122143; AAM52655.1; -; mRNA.
DR   FlyBase; FBgn0040765; luna.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 3.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 3.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW   Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ   SEQUENCE   570 AA;  61770 MW;  F889881AFBC85F3D CRC64;

Alignment Scores:
Pred. No.:              11.7          Length:         570
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q9V5S5_DROME (1-570)

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||||
Db        357 SerSerSerSerHisSerSerSerSerSer 366

RESULT 38
Q8BZY1_MOUSE
ID   Q8BZY1_MOUSE PRELIMINARY;      PRT;   578 AA.
AC   Q8BZY1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mus musculus 15 days embryo male testis cDNA, RIKEN full-length
DE   enriched library, clone:8030474K03 product:hypothetical Type I
DE   antifreeze protein containing protein, full insert sequence.
DE   (Fragment).
GN   Name=8030474K03Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
```

```
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK033245; BAC28209.1; -; mRNA.
DR   Ensembl; ENSMUSG00000046774; Mus musculus.
DR   MGI; MGI:2685988; 8030474K03Rik.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   578 AA;  60998 MW;  195078DBCBB48BB9 CRC64;

Alignment Scores:
Pred. No.:              11.7          Length:        578
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0
```

```
US-10-677-865A-2 (1-1590) x Q8BZY1_MOUSE (1-578)

Qy        612 TACCCCAGAGGTTTGGGAAGAGGAAGAGCT 583
              ||||||||||||||||||||||||||||||
Db        220 TyrProArgGlyLeuGlyArgGlyArgAla 229

RESULT 39
Q96E25_HUMAN
ID   Q96E25_HUMAN PRELIMINARY;      PRT;   609 AA.
AC   Q96E25;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   ZNF503 protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Skin, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC013011; AAH13011.1; -; mRNA.
DR   EMBL; BC021165; AAH21165.1; -; mRNA.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ   SEQUENCE   609 AA;  59292 MW;  4A1D2536F82CCFD7 CRC64;

Alignment Scores:
Pred. No.:              11.6          Length:        609
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0
```

```
US-10-677-865A-2 (1-1590) x Q96E25_HUMAN (1-609)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        275 GlyGlySerSerGlySerSerSerGlySer 284

RESULT 40
Q54MS1_DICDI
ID   Q54MS1_DICDI PRELIMINARY;      PRT;   612 AA.
AC   Q54MS1;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=DDB0186665;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAFI01000124; EAL64687.1; -; Genomic_DNA.
DR   InterPro; IPR004861; Put_tyr_phos.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   Pfam; PF03162; Y_phosphatase2; 1.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
KW   Hypothetical protein; Nucleotide-binding.
SQ   SEQUENCE   612 AA;  67336 MW;  D19196E3C60BA00A CRC64;

Alignment Scores:
Pred. No.:              11.6          Length:        612
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q54MS1_DICDI (1-612)

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCC 237
              ||||||||||||||||||||||||||||||
Db        518 GlySerSerSerSerGlySerSerIleSer 527

RESULT 41
Q7PFH8_ANOGA
ID   Q7PFH8_ANOGA PRELIMINARY;      PRT;   642 AA.
AC   Q7PFH8;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ENSANGP00000022995 (Fragment).
GN   ORFNames=ENSANGG00000016503;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC   Anophelinae; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PEST;
RG   The Anopheles gambiae Sequence Committee;
RT   "Anopheles gambiae re-annotation.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PEST;
RG   The Anopheles gambiae Sequence Committee;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008823; EAA45310.2; -; Genomic_DNA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   InterPro; IPR012725; DnaK_prok.
DR   InterPro; IPR002048; EF_hand_Ca_bd.
DR   InterPro; IPR001023; Hsp70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 1.
DR   TIGRFAMs; TIGR02350; prok_dnaK; 1.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR   PROSITE; PS00297; HSP70_1; 1.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   ATP-binding; Nucleotide-binding.
FT   NON_TER       1      1
SQ   SEQUENCE   642 AA;  69443 MW;  E3B8CDF6DACADE9B CRC64;

Alignment Scores:
Pred. No.:              11.6          Length:        642
Score:                  10.00         Matches:       10
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.9%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q7PFH8_ANOGA (1-642)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        618 GlySerSerGlySerSerSerGlySerSer 627

RESULT 42
Q96F45_HUMAN
ID   Q96F45_HUMAN PRELIMINARY;      PRT;   646 AA.
AC   Q96F45;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Zinc finger protein 503.
GN   Name=ZNF503;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterus;
RA   Director MGC Project;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC011625; AAH11625.1; -; mRNA.
DR   Ensembl; ENSG00000165655; Homo sapiens.
DR   HGNC; HGNC:23589; ZNF503.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ   SEQUENCE   646 AA;  62555 MW;  F7F341E6C8170B85 CRC64;

Alignment Scores:
Pred. No.:              11.6          Length:         646
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q96F45_HUMAN (1-646)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        312 GlyGlySerSerGlySerSerSerGlySer 321

RESULT 43
GSPD2_ERWCH
ID   GSPD2_ERWCH    STANDARD;      PRT;   710 AA.
AC   Q01565;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   General secretion pathway protein D precursor (Pectic enzymes
DE   secretion protein outD).
GN   Name=outD;
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Dickeya.
OX   NCBI_TaxID=556;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   STRAIN=3937;
RX   MEDLINE=93086427; PubMed=1453958;
RA   Condemine G., Dorel C., Hugouvieux-Cotte-Pattat N., Robert-Baudouy J.;
RT   "Some of the out genes involved in the secretion of pectate lyases in
RT   Erwinia chrysanthemi are regulated by kdgR.";
RL   Mol. Microbiol. 6:3199-3211(1992).
CC   -!- FUNCTION: Involved in a general secretion pathway (GSP) for the
CC       export of proteins. Required for the translocation of the multiple
CC       pectic enzymes.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (Probable).
CC   -!- SIMILARITY: Belongs to the GSP D family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; X65265; CAA46370.1; -; Genomic_DNA.
DR   PIR; S28014; S28014.
DR   InterPro; IPR001775; Bac_GSPD.
DR   InterPro; IPR004846; GSPII/IIIprotein.
DR   InterPro; IPR004845; GSPIIproteinC.
DR   InterPro; IPR005644; NolW-like.
DR   Pfam; PF00263; Secretin; 1.
DR   Pfam; PF03958; Secretin_N; 3.
DR   PRINTS; PR00811; BCTERIALGSPD.
DR   PROSITE; PS00875; T2SP_D; 1.
KW   Membrane; Outer membrane; Signal; Transport.
FT   SIGNAL        1     27       Potential.
FT   CHAIN        28    710       General secretion pathway protein D.
FT   COMPBIAS    288    353       Gly/Ser-rich.
SQ   SEQUENCE   710 AA;  76214 MW;  156E84CC50CD54FA CRC64;

Alignment Scores:
Pred. No.:              11.5          Length:         710
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x GSPD2_ERWCH (1-710)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        321 GlySerSerGlySerSerSerGlySerSer 330

RESULT 44
Q6BIY3_DEBHA
ID   Q6BIY3_DEBHA PRELIMINARY;      PRT;   770 AA.
AC   Q6BIY3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Debaryomyces hansenii chromosome G of strain CBS767 of Debaryomyces
DE   hansenii.
GN   OrderedLocusNames=DEHA0G07337g;
OS   Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX   NCBI_TaxID=4959;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=ATCC 36239 / CBS 767;
RX   PubMed=15229592; DOI=10.1038/nature02579;
RA   Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA   Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA   Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA   Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA   Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA   Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA   Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA   Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA   Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA   Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA   Swennen D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA   Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA   Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA   Wincker P., Souciet J.-L.;
```

```
RT   "Genome evolution in yeasts.";
RL   Nature 430:35-44(2004).
DR   EMBL; CR382139; CAG90299.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   770 AA;  76049 MW;  F1ABCD7AC396E1D1 CRC64;

Alignment Scores:
Pred. No.:              11.4          Length:         770
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q6BIY3_DEBHA (1-770)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        599 GlySerSerGlySerSerSerGlySerSer 608

RESULT 45
Q4QG78_LEIMA
ID   Q4QG78_LEIMA PRELIMINARY;      PRT;  2063 AA.
AC   Q4QG78;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=LmjF13.0760;
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Friedlin;
RA   Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA   Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,
RA   Saunders D., Seeger K., Warren T., Apostolou Z., Bauser C., Beck A.,
RA   Bianchettin G., Borzym K., Bothe G., Bruschi C., Ciarloni L.,
RA   Duesterhoeft A., Fuchs M., Gabel C., Goffeau A., Hilbert H.,
RA   Klages S., Kube M., Marra M., Masuy D., Mueller-Auer S., Pohl T.,
RA   Purnelle B., Reinhardt R., Rieger M., Robben J., Schaefer M.,
RA   Tosato V., Volckaert G., Wambutt R., Wedler H., Zimmermann M.,
RA   Rajandream M., and Barrell B.G.;
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CT005252; CAJ02713.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   2063 AA;  218311 MW;  7E81446A6E0976C0 CRC64;

Alignment Scores:
Pred. No.:              10.3          Length:         2063
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q4QG78_LEIMA (1-2063)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db       1332 SerSerSerHisSerSerSerSerSerSer 1341

RESULT 46
SPEN_DROME
ID   SPEN_DROME     STANDARD;      PRT;  5560 AA.
AC   Q8SX83; Q9NHN1; Q9NJ17; Q9U6C3; Q9VPL1; Q9VPL2;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Split ends protein.
GN   Name=spen; ORFNames=CG18497;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORMS 3 AND 4), FUNCTION, ALTERNATIVE PROMOTER
RP   USAGE, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND MUTANTS E9 AND
RP   D57.
RC   TISSUE=Embryo;
RX   MEDLINE=20025936; PubMed=10556062;
RA   Wiellette E.L., Harding K.W., Mace K.A., Ronshaugen M.R., Wang F.Y.,
RA   McGinnis W.;
RT   "spen encodes an RNP motif protein that interacts with Hox pathways to
RT   repress the development of head-like sclerites in the Drosophila
RT   trunk.";
RL   Development 126:5373-5385(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 4).
RC   TISSUE=Embryo;
RX   MEDLINE=20157049; PubMed=10655223;
RA   Rebay I., Chen F., Hsiao F., Kolodziej P.A., Kuang B.H., Laverty T.,
RA   Suh C., Voas M., Williams A., Rubin G.M.;
RT   "A genetic screen for novel components of the Ras/mitogen-activated
RT   protein kinase signaling pathway that interact with the yan gene of
RT   Drosophila identifies split ends, a new RNA recognition motif-
RT   containing protein.";
RL   Genetics 154:695-712(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Embryo;
RX   MEDLINE=20171275; PubMed=10704397;
RA   Kuang B.H., Wu S.C.-Y., Shin Y.-A., Luo L., Kolodziej P.A.;
RT   "split ends encodes large nuclear proteins that regulate neuronal cell
RT   fate and axon extension in the Drosophila embryo.";
RL   Development 127:1517-1529(2000).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND TISSUE SPECIFICITY.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
```

```
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [5]
RP   GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [6]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 424-2002.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [7]
RP   FUNCTION.
RX   MEDLINE=20253107; PubMed=10790398;
RA   Lane M.E., Elend M., Heidmann D., Herr A., Marzodko S., Herzig A.,
RA   Lehner C.F.;
RT   "A screen for modifiers of cyclin E function in Drosophila
RT   melanogaster identifies Cdk2 mutations, revealing the insignificance
RT   of putative phosphorylation sites in Cdk2.";
RL   Genetics 155:233-244(2000).
RN   [8]
RP   FUNCTION ON EGF RECEPTOR PATHWAY.
RX   MEDLINE=20414403; PubMed=10959845; DOI=10.1016/S0960-9822(00)00625-4;
RA   Chen F., Rebay I.;
RT   "split ends, a new component of the Drosophila EGF receptor pathway,
RT   regulates development of midline glial cells.";
RL   Curr. Biol. 10:943-946(2000).
RN   [9]
RP   FUNCTION ON WG PATHWAY.
RX   MEDLINE=22668876; PubMed=12783785; DOI=10.1242/dev.00527;
RA   Lin H.V., Doroquez D.B., Cho S., Chen F., Rebay I., Cadigan K.M.;
RT   "Splits ends is a tissue/promoter specific regulator of Wingless
RT   signaling.";
RL   Development 130:3125-3135(2003).
CC   -!- FUNCTION: Probable corepressor protein, which regulates different
CC       key pathways such as the EGF receptor and Wg pathways. Involved in
CC       neuronal cell fate, survival and axon guidance, cell cycle
CC       regulation and repression of head identity in the embryonic trunk.
CC       May act with the Hox gene Deformed and the EGF receptor signaling
CC       pathway. Positive regulator of the Wg pathway in larval tissues
CC       but not in embryonic tissues. May act as a transcriptional
CC       corepressor protein, which repress transcription via the
CC       recruitment of large complexes containing histone deacetylase
CC       proteins.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative promoter;
CC         Comment=2 isoforms, 1 (shown here) and 2, are produced by use of
CC         alternative promoters;
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1;
CC         IsoId=Q8SX83-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q8SX83-2; Sequence=VSP_008565, VSP_008566;
CC         Note=No experimental confirmation available;
CC       Name=3; Synonyms=SpenL;
CC         IsoId=Q8SX83-3; Sequence=VSP_008567;
CC         Note=Produced by alternative splicing of isoform 1;
CC       Name=4; Synonyms=SpenS;
CC         IsoId=Q8SX83-4; Sequence=VSP_008565, VSP_008566, VSP_008567;
CC         Note=Produced by alternative splicing of isoform 2;
CC   -!- TISSUE SPECIFICITY: Ubiquitous. Expressed prior to cellularization
CC       in stage 3 embryos, and in blastoderm cells, including pole cells.
CC       Expressed throughout the rest of embryogenesis. Later, it is
CC       expressed at higher level in epidermal cells and CNS.
CC   -!- DEVELOPMENTAL STAGE: Isoform 3 is expressed both maternally and
CC       zygotically.
CC   -!- SIMILARITY: Belongs to the RNP Spen family.
CC   -!- SIMILARITY: Contains 3 RRM (RNA recognition motif) domains.
CC   -!- SIMILARITY: Contains 1 SPOC domain.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF188205; AAF13218.1; -; mRNA.
DR   EMBL; AF184612; AAF26299.1; -; mRNA.
DR   EMBL; AF221715; AAF34661.1; ALT_INIT; mRNA.
DR   EMBL; AE003590; AAF51534.2; -; Genomic_DNA.
DR   EMBL; AE003590; AAF51535.2; -; Genomic_DNA.
DR   EMBL; AE003590; AAN10511.1; -; Genomic_DNA.
DR   EMBL; AY094788; AAM11141.1; ALT_SEQ; mRNA.
DR   HSSP; P09651; 1HA1.
DR   Ensembl; CG18497; Drosophila melanogaster.
DR   FlyBase; FBgn0016977; spen.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0030528; F:transcription regulator activity; IMP.
DR   GO; GO:0007411; P:axon guidance; IMP.
DR   GO; GO:0007173; P:epidermal growth factor receptor signaling . . .; IMP.
DR   GO; GO:0007403; P:glial cell fate determination; IMP.
DR   GO; GO:0007400; P:neuroblast fate determination; IMP.
DR   GO; GO:0045449; P:regulation of transcription; IMP.
DR   GO; GO:0007379; P:segment specification; IMP.
DR   GO; GO:0016055; P:Wnt receptor signaling pathway; IMP.
DR   InterPro; IPR012677; a_b_plait_nuc_bd.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR000354; Involucrin_rpt.
DR   InterPro; IPR000504; RNP1_RNA_bd.
DR   InterPro; IPR012921; SPOC.
DR   InterPro; IPR010912; SPOC_met.

Alignment Scores:
Pred. No.:                9.41          Length:          5560
Score:                    10.00         Matches:         10
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.9%          Indels:          0
DB:                       1             Gaps:            0

US-10-677-865A-2 (1-1590) x SPEN_DROME (1-5560)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 47
Q7UK51_RHOBA
ID   Q7UK51_RHOBA PRELIMINARY;      PRT;    46 AA.
AC   Q7UK51;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RB10861;
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
```

```
OX   NCBI_TaxID=117;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294152; CAD77030.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   46 AA;  4865 MW;  940E7EF225285141 CRC64;

Alignment Scores:
Pred. No.:              127           Length:        46
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q7UK51_RHOBA (1-46)

Qy        179 GTGACTCCAGCGGCTTCAGTAGCTACA 205
              ||||||||||||||||||||||||||
Db         12 ValThrProAlaAlaSerValAlaThr 20

RESULT 48
Q4YEV7_PLABE
ID   Q4YEV7_PLABE PRELIMINARY;      PRT;    85 AA.
AC   Q4YEV7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PB404986.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01005830; CAI03462.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
FT   NON_TER      85     85
SQ   SEQUENCE   85 AA;  9656 MW;  A37D9A72A3E643A1 CRC64;

Alignment Scores:
Pred. No.:              120           Length:        85
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q4YEV7_PLABE (1-85)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||
Db          8 SerSerSerHisSerSerSerSerSer 16

RESULT 49
Q5BQQ1_SCHJA
ID   Q5BQQ1_SCHJA PRELIMINARY;      PRT;    92 AA.
AC   Q5BQQ1;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY915915; AAX31136.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   92 AA;  9296 MW;  1E29198FA74BEAE8 CRC64;

Alignment Scores:
Pred. No.:              119           Length:        92
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5BQQ1_SCHJA (1-92)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db         23 SerSerGlySerSerSerGlySerSer 31

RESULT 50
Q71D57_DROYA
ID   Q71D57_DROYA PRELIMINARY;      PRT;   109 AA.
AC   Q71D57;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CG15188 (Fragment).
OS   Drosophila yakuba (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7245;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22887302; PubMed=14525923; DOI=10.1101/gr.1311003;
RA   Domazet-Loso T., Tautz D.;
RT   "An evolutionary analysis of orphan genes in Drosophila.";
RL   Genome Res. 13:2213-2219(2003).
DR   EMBL; AF532013; AAQ09912.1; -; mRNA.
FT   NON_TER       1      1
FT   NON_TER     109    109
SQ   SEQUENCE   109 AA;  11347 MW;  D2DF20FD94E83C42 CRC64;

Alignment Scores:
Pred. No.:              117           Length:        109
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q71D57_DROYA (1-109)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db         91 SerSerGlySerSerSerGlySerSer 99
```

```
RESULT 51
Q7U8W2_SYNPX
ID   Q7U8W2_SYNPX PRELIMINARY;      PRT;   109 AA.
AC   Q7U8W2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ferredoxin.
GN   Name=petF1; OrderedLocusNames=SYNW0497;
OS   Synechococcus sp. (strain WH8102).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=84588;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA   Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA   Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA   Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT   "The genome of a motile marine Synechococcus.";
RL   Nature 424:1037-1042(2003).
DR   EMBL; BX569690; CAE07012.1; -; Genomic_DNA.
DR   HSSP; P27320; 1DOX.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0005506; F:iron ion binding; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001041; Ferredoxin.
DR   Pfam; PF00111; Fer2; 1.
KW   Complete proteome; Iron; Iron-sulfur; Metal-binding.
SQ   SEQUENCE   109 AA;  11507 MW;  359AD7BBBD155AA9 CRC64;

Alignment Scores:
Pred. No.:              117          Length:        109
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x Q7U8W2_SYNPX (1-109)

Qy        195 GAAGCCGCTGGAGTCACCCTTCCCAGT 169
              |||||||||||||||||||||||||||
Db         33 GluAlaAlaGlyValThrLeuProSer 41

RESULT 52
Q63UZ0_BURPS
ID   Q63UZ0_BURPS PRELIMINARY;      PRT;   141 AA.
AC   Q63UZ0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Putative cytochrome C oxidase.
GN   OrderedLocusNames=BPSL1454;
OS   Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Burkholderia; pseudomallei group.
OX   NCBI_TaxID=28450;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=K96243;
RX   PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA   Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA   Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA   Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,
RA   Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA   Chillingworth T., Cronin A., Crossett B., Davis P., DeShazer D.,
RA   Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA   Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA   Rabbinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA   Songsivilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA   Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT   "Genomic plasticity of the causative agent of melioidosis,
RT   Burkholderia pseudomallei.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR   EMBL; BX571965; CAH35456.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001505; Copper_CuA.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
DR   ProDom; PD000131; Copper_CuA; 1.
DR   PROSITE; PS50857; COX2_CUA; 1.
KW   Complete proteome.
SQ   SEQUENCE   141 AA;  15096 MW;  86CF851849FF1773 CRC64;

Alignment Scores:
Pred. No.:              114          Length:        141
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x Q63UZ0_BURPS (1-141)

Qy        429 GCTTCCCGAGTGAGAGCTGCTGCTCCC 403
              |||||||||||||||||||||||||||
Db         44 AlaSerArgValArgAlaAlaAlaPro 52

RESULT 53
Q62JQ5_BURMA
ID   Q62JQ5_BURMA PRELIMINARY;      PRT;   141 AA.
AC   Q62JQ5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Cytochrome c oxidase family protein.
GN   OrderedLocusNames=BMA1408;
OS   Burkholderia mallei (Pseudomonas mallei).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Burkholderia.
OX   NCBI_TaxID=13373;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 23344;
RX   PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA   Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA   Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA   Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA   Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA   Madupu R., Mohammoud Y., Nelson W.C., Radune D., Romero C.M.,
RA   Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA   Zafar N., Zhou L., Fraser C.M.;
RT   "Structural flexibility in the Burkholderia mallei genome.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR   EMBL; CP000010; AAU47609.1; -; Genomic_DNA.
DR   TIGR; BMA1408; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001505; Copper_CuA.
DR   InterPro; IPR002429; Cyt_c_ox_2.
DR   Pfam; PF00116; COX2; 1.
DR   ProDom; PD000131; Copper_CuA; 1.
DR   PROSITE; PS50857; COX2_CUA; 1.
KW   Complete proteome.
SQ   SEQUENCE   141 AA;  15126 MW;  EDA228AE99BE0322 CRC64;

Alignment Scores:
Pred. No.:              114          Length:        141
Score:                  9.00         Matches:       9
```

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x Q62JQ5_BURMA (1-141)

Qy        429 GCTTCCCGAGTGAGAGCTGCTGCTCCC 403
                ||||||||||||||||||||||||||
Db         44 AlaSerArgValArgAlaAlaAlaPro 52

RESULT 54
Q4TL72_9SPHN
ID   Q4TL72_9SPHN PRELIMINARY;      PRT;   152 AA.
AC   Q4TL72;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Acetyl-CoA carboxylase.
GN   ORFNames=ELI2405;
OS   Erythrobacter litoralis HTCC2594.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Erythrobacter.
OX   NCBI_TaxID=314225;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HTCC2594;
RA   Giovannoni S.J., Cho J.-C., Ferriera S., Johnson J., Kravitz S.,
RA   Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
RA   Friedman R., Venter J.C.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAGG01000008; EAL74078.1; -; Genomic_DNA.
SQ   SEQUENCE   152 AA;  15358 MW;  7ED39FCDEF6DA254 CRC64;

Alignment Scores:
Pred. No.:              114             Length:        152
Score:                  9.00            Matches:       9
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x Q4TL72_9SPHN (1-152)

Qy       1280 GCACCGGCTCCTTCAGCAGCAGCTCCA 1306
                ||||||||||||||||||||||||||
Db         56 AlaProAlaProSerAlaAlaAlaPro 64

RESULT 55
Q5CQE7_CRYPV
ID   Q5CQE7_CRYPV PRELIMINARY;      PRT;   153 AA.
AC   Q5CQE7;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=cgd4_1000;
OS   Cryptosporidium parvum.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC   Cryptosporidiidae; Cryptosporidium.
OX   NCBI_TaxID=5807;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Iowa type II;
RX   PubMed=15044751; DOI=10.1126/science.1094786;
RA   Abrahamsen M.S., Templeton T.J., Enomoto S., Abrahante J.E., Zhu G.,
RA   Lancto C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA   Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA   Anantharaman V., Aravind L., Kapur V.;
RT   "Complete genome sequence of the apicomplexan, Cryptosporidium
RT   parvum.";
RL   Science 304:441-445(2004).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAEE01000009; EAK87639.1; -; Genomic_DNA.
DR   InterPro; IPR006209; EGF_like.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   153 AA;  15283 MW;  3C69BB93A35709F0 CRC64;

Alignment Scores:
Pred. No.:              113             Length:        153
Score:                  9.00            Matches:       9
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x Q5CQE7_CRYPV (1-153)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
                ||||||||||||||||||||||||||
Db         50 SerGlySerSerSerSerGlySerSer 58

RESULT 56
Q4SGP7_TETNG
ID   Q4SGP7_TETNG PRELIMINARY;      PRT;   155 AA.
AC   Q4SGP7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 3 SCAF14593, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00018534001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014593; CAG00185.1; -; Genomic_DNA.
FT   NON_TER       1      1
SQ   SEQUENCE   155 AA;  16605 MW;  2248D18E75E438CE CRC64;

Alignment Scores:
Pred. No.:              113             Length:        155
Score:                  9.00            Matches:       9
Percent Similarity:     100.0%          Conservative:  0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.7%            Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q4SGP7_TETNG (1-155)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db         64 SerSerGlySerSerSerGlySerSer 72

RESULT 57
Q9EWB3_STRLI
ID   Q9EWB3_STRLI PRELIMINARY;      PRT;   158 AA.
AC   Q9EWB3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Twin arginine translocation complex component.
GN   Name=tatB;
OS   Streptomyces lividans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1916;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21555085; PubMed=11698358;
RX   DOI=10.1128/JB.183.23.6727-6732.2001;
RA   Schaerlaekens K.M., Schierova M., Lammertyn E., Geukens N., Anne J.,
RA   Van Mellaert L.;
RT   "Twin-arginine translocation pathway in Streptomyces lividans.";
RL   J. Bacteriol. 183:6727-6732(2001).
DR   EMBL; AJ292256; CAC21724.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0015450; F:protein translocase activity; IEA.
DR   GO; GO:0009306; P:protein secretion; IEA.
DR   GO; GO:0006605; P:protein targeting; IEA.
DR   GO; GO:0015031; P:protein transport; IEA.
DR   InterPro; IPR003998; TatB.
DR   PRINTS; PR01506; TATBPROTEIN.
SQ   SEQUENCE   158 AA;  17402 MW;  FC704E6558DB4F14 CRC64;

Alignment Scores:
Pred. No.:              113             Length:         158
Score:                  9.00            Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.7%            Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x Q9EWB3_STRLI (1-158)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        113 SerSerGlySerSerSerGlySerSer 121

RESULT 58
TATB_STRCO
ID   TATB_STRCO     STANDARD;      PRT;   161 AA.
AC   Q9FBK8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Sec-independent protein translocase protein tatB homolog.
GN   Name=tatB; OrderedLocusNames=SCO5150; ORFNames=SCP8.13;
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- FUNCTION: Required for correct localization of precursor proteins
CC       bearing signal peptides with the twin arginine conserved motif
CC       S/T-R-R-X-F-L-K. This sec-independent pathway is termed TAT for
CC       twin-arginine translocation system. This system mainly transports
CC       proteins with bound cofactors that require folding prior to export
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC   -!- SIMILARITY: Belongs to the tatB family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AL939122; CAC01351.1; -; Genomic_DNA.
DR   HAMAP; MF_00237; -; 1.
DR   InterPro; IPR003998; TatB.
DR   PRINTS; PR01506; TATBPROTEIN.
KW   Complete proteome; Membrane; Protein transport; Translocation;
KW   Transmembrane; Transport.
FT   TRANSMEM      2     22       Potential.
SQ   SEQUENCE   161 AA;  17757 MW;  0510C0D221BEEEF3 CRC64;

Alignment Scores:
Pred. No.:              113             Length:         161
Score:                  9.00            Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.7%            Indels:         0
DB:                     1               Gaps:           0

US-10-677-865A-2 (1-1590) x TATB_STRCO (1-161)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        113 SerSerGlySerSerSerGlySerSer 121

RESULT 59
Q5T4F5_HUMAN
ID   Q5T4F5_HUMAN PRELIMINARY;      PRT;   167 AA.
AC   Q5T4F5;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Novel protein.
GN   Name=C10orf132; ORFNames=RP11-459F3.4-001;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Almeida J.;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL358938; CAI14276.1; -; Genomic_DNA.
DR   Ensembl; ENSG00000155265; Homo sapiens.
DR   HGNC; HGNC:31668; C10orf132.
SQ   SEQUENCE   167 AA;  18334 MW;  1DB5D42080A4A66C CRC64;
```

```
Alignment Scores:
Pred. No.:              112           Length:        167
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5T4F5_HUMAN (1-167)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db        143 SerGlySerSerSerSerGlySerSer 151

RESULT 60
Q9M176_ARATH
ID   Q9M176_ARATH PRELIMINARY;      PRT;   167 AA.
AC   Q9M176;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein T5C2_130 (Hypothetical protein At3g43430).
GN   Name=T5C2_130; Synonyms=At3g43430;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA   Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA   Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA   Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; AL138664; CAB81584.1; -; Genomic_DNA.
DR   EMBL; AK175367; BAD43130.1; -; mRNA.
DR   PIR; T47385; T47385.
DR   HSSP; Q9LRB7; 1IYM.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   167 AA;  19001 MW;  B098D87CC2B3DEC2 CRC64;

Alignment Scores:
Pred. No.:              112           Length:        167
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q9M176_ARATH (1-167)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||
Db         41 SerSerSerSerSerSerGlnSerSer 49

RESULT 61
Q9D428_MOUSE
ID   Q9D428_MOUSE PRELIMINARY;      PRT;   167 AA.
AC   Q9D428;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE   library, clone:4933417O08 product:similar to SIMILAR TO HYPOTHETICAL
DE   PROTEIN, MNCB-1213 (RIKEN cDNA 4933417O08).
GN   Name=4933417O08Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
```

```
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Director MGC Project;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK016850; BAB30463.1; -; mRNA.
DR   EMBL; BC049116; AAH49116.1; -; mRNA.
DR   Ensembl; ENSMUSG00000042532; Mus musculus.
DR   MGI; MGI:1918396; 4933417O08Rik.
KW   Hypothetical protein.
SQ   SEQUENCE   167 AA;  18335 MW;  4A34CB2080A4B167 CRC64;

Alignment Scores:
Pred. No.:              112           Length:         167
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q9D428_MOUSE (1-167)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        143 SerGlySerSerSerSerGlySerSer 151

RESULT 62
Q69NY7_ORYSA
ID   Q69NY7_ORYSA PRELIMINARY;      PRT;   185 AA.
AC   Q69NY7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein OSJNBb0050B07.2.
GN   Name=OSJNBb0050B07.2;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OSJNBb0050B07.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005632; BAD31580.1; -; Genomic_DNA.
DR   Gramene; Q69NY7; -.
DR   InterPro; IPR010399; Zim.
DR   Pfam; PF06200; Zim; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   185 AA;  19486 MW;  DB7B83C47CE5ED73 CRC64;

Alignment Scores:
Pred. No.:              111           Length:         185
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q69NY7_ORYSA (1-185)

Qy        403 GGGAGCAGCAGCTCTCACTCGGGAAGC 429
              ||||||||||||||||||||||||||
Db        114 GlySerSerSerSerHisSerGlySer 122

RESULT 63
Q5LR68_SILPO
ID   Q5LR68_SILPO PRELIMINARY;      PRT;   185 AA.
AC   Q5LR68;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
```

```
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Phage prohead protease, HK97 family.
GN   OrderedLocusNames=SPO2262;
OS   Silicibacter pomeroyi.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC   Rhodobacteraceae; Silicibacter.
OX   NCBI_TaxID=89184;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX   PubMed=15602564; DOI=10.1038/nature03170;
RA   Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA   Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA   Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA   Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA   Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA   Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA   Haft D.H., Selengut J., Ward N.;
RT   "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT   marine environment.";
RL   Nature 432:910-913(2004).
DR   EMBL; CP000031; AAV95527.1; -; Genomic_DNA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   InterPro; IPR006433; Peptidase_U35.
DR   Pfam; PF04586; Peptidase_U35; 1.
DR   TIGRFAMs; TIGR01543; proheadase_HK97; 1.
KW   Complete proteome; Protease.
SQ   SEQUENCE   185 AA;  20346 MW;  78D490C703E5800F CRC64;

Alignment Scores:
Pred. No.:              111           Length:        185
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5LR68_SILPO (1-185)

Qy        443 CGGGAAGCAGCAGCTCTCATTCGAGCA 469
              |||||||||||||||||||||||||||
Db        101 ArgGluAlaAlaAlaLeuIleArgAla 109

RESULT 64
Q90YE0_XENLA
ID   Q90YE0_XENLA PRELIMINARY;      PRT;   185 AA.
AC   Q90YE0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Homeobox protein Hox A11 (Fragment).
GN   Name=hoxa11;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21366019; PubMed=11472855; DOI=10.1016/S0925-4773(01)00438-5;
RA   Lombardo A., Slack J.M.W.;
RT   "Abdominal B-type Hox gene activity in Xenopus laevis.";
RL   Mech. Dev. 106:191-195(2001).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR   EMBL; AJ319668; CAC44977.1; -; mRNA.
DR   HSSP; P09631; 1PUF.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; Homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   185 AA;  21263 MW;  704FC07AA4A70EA8 CRC64;

Alignment Scores:
Pred. No.:              111           Length:        185
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q90YE0_XENLA (1-185)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||
Db         89 SerSerSerSerSerSerGlnSerSer 97

RESULT 65
Q7ZZY5_CHICK
ID   Q7ZZY5_CHICK PRELIMINARY;      PRT;   186 AA.
AC   Q7ZZY5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Decoy receptor 3 (Fragment).
GN   Name=DcR3;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22760275; PubMed=12878204; DOI=10.1016/S0006-291X(03)01304-4;
RA   Bridgham J.T., Johnson A.L.;
RT   "Characterization of chicken TNFR superfamily decoy receptors, DcR3
RT   and osteoprotegerin.";
RL   Biochem. Biophys. Res. Commun. 307:956-961(2003).
DR   EMBL; AY251406; AAP03889.1; -; mRNA.
DR   HSSP; O14763; 1DU3.
DR   Ensembl; ENSGALG00000006106; Gallus gallus.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00020; TNFR_c6; 2.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR   PROSITE; PS00652; TNFR_NGFR_1; 1.
DR   PROSITE; PS50050; TNFR_NGFR_2; 1.
KW   Receptor.
FT   NON_TER     186    186
SQ   SEQUENCE   186 AA;  20671 MW;  31D65731DACB758E CRC64;

Alignment Scores:
Pred. No.:              111           Length:        186
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q7ZZY5_CHICK (1-186)

Qy       1285 GGCTCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||||||
Db        155 GlySerPheSerSerSerSerSerSer 163
```

```
RESULT 66
Q7YY80_CRYPV
ID   Q7YY80_CRYPV PRELIMINARY;      PRT;   192 AA.
AC   Q7YY80;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serine repeat antigen, probable (Secreted protein with 3
DE   cryptosporidium-specific paralogs, serine stretch at C terminus,
DE   possible mucin, transcripts identified by EST).
GN   ORFNames=56k.04, cgd6_40;
OS   Cryptosporidium parvum.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC   Cryptosporidiidae; Cryptosporidium.
OX   NCBI_TaxID=5807;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Iowa;
RA   Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,
RA   Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
RT   "Integrated mapping, chromosomal sequencing and sequence analysis of
RT   Cryptosporidium parvum.";
RL   Genome Res. 0:0-0(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Iowa type II;
RX   PubMed=15044751; DOI=10.1126/science.1094786;
RA   Abrahamsen M.S., Templeton T.J., Enomoto S., Abrahante J.E., Zhu G.,
RA   Lancto C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA   Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA   Anantharaman V., Aravind L., Kapur V.;
RT   "Complete genome sequence of the apicomplexan, Cryptosporidium
RT   parvum.";
RL   Science 304:441-445(2004).
DR   EMBL; BX538350; CAD98606.1; -; Genomic_DNA.
DR   EMBL; AAEE01000014; EAK87333.1; -; Genomic_DNA.
SQ   SEQUENCE   192 AA;  19686 MW;  CB363BF9E9B40F77 CRC64;

Alignment Scores:
Pred. No.:              111           Length:        192
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q7YY80_CRYPV (1-192)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db        158 SerGlySerSerSerSerGlySerSer 166

RESULT 67
Q5CG28_CRYHO
ID   Q5CG28_CRYHO PRELIMINARY;      PRT;   192 AA.
AC   Q5CG28;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Serine repeat antigen.
GN   ORFNames=Chro.60014;
OS   Cryptosporidium hominis.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC   Cryptosporidiidae; Cryptosporidium.
OX   NCBI_TaxID=237895;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=TU502;
RA   Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G.,
RA   Puiu D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H.,
RA   Bankier A.T., Peterson D.L., Abrahamsen M.S., Kapur V., Tzipori S.,
RA   Buck G.A.;
RT   "The genome of Cryptosporidium hominis.";
RL   Nature 431:1107-1112(2004).
DR   EMBL; AAEL01000376; EAL35559.1; -; Genomic_DNA.
SQ   SEQUENCE   192 AA;  19936 MW;  CA506BCCC66E5E41 CRC64;

Alignment Scores:
Pred. No.:              111           Length:        192
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5CG28_CRYHO (1-192)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db        169 SerGlySerSerSerSerGlySerSer 177

RESULT 68
Q8XUT1_RALSO
ID   Q8XUT1_RALSO PRELIMINARY;      PRT;   194 AA.
AC   Q8XUT1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   PUTATIVE CALCIUM BINDING SIGNAL PEPTIDE PROTEIN.
GN   OrderedLocusNames=RSc3104; ORFNames=RS00557;
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646073; CAD16813.1; -; Genomic_DNA.
DR   HSSP; P38505; 1JFJ.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   InterPro; IPR011992; EF-Hand_type.
DR   InterPro; IPR002048; EF_hand_Ca_bd.
DR   Pfam; PF00036; efhand; 3.
DR   ProDom; PD000012; EF-hand; 1.
DR   SMART; SM00054; EFh; 3.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_3.
KW   Calcium; Complete proteome; Repeat.
SQ   SEQUENCE   194 AA;  20809 MW;  AFAA640E3D0E0F0F CRC64;

Alignment Scores:
Pred. No.:              111           Length:        194
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q8XUT1_RALSO (1-194)

Qy       1280 GCACCGGCTCCTTCAGCAGCAGCTCCA 1306
              ||||||||||||||||||||||||||
Db         37 AlaProAlaProSerAlaAlaAlaPro 45

RESULT 69
Q94476_DICDI
ID   Q94476_DICDI PRELIMINARY;      PRT;   199 AA.
```

```
AC   Q94476;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ORFveg113 (Fragment).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Loomis W.F.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U66523; AAB06786.1; -; mRNA.
DR   HSSP; P11376; 2EQL.
DR   DictyBase; DDB0185118; rcdBB.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   InterPro; IPR001916; Glyco_hydro_22.
DR   InterPro; IPR008258; LT_catalytic.
DR   Pfam; PF01464; SLT; 1.
DR   SMART; SM00263; LYZ1; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   199 AA;  18935 MW;  DBE0D17B1BC81A3A CRC64;

Alignment Scores:
Pred. No.:              111           Length:         199
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q94476_DICDI (1-199)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        143 SerSerGlySerSerSerGlySerSer 151

RESULT 70
Q8MML4_DICDI
ID   Q8MML4_DICDI PRELIMINARY;      PRT;   202 AA.
AC   Q8MML4; Q8T171;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   HypotPutative Glycosidase (Similar to Dictyostelium discoideum (Slime
DE   mold). ORFveg113) (Random cDNA clone veg113).
GN   Name=rcdBB; ORFNames=DDB0185118;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA   Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA   Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RX   MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA   Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA   Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA   Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA   Platzer M., Rosenthal A., Noegel A.A.;
RT   "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL   Nature 418:79-85(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Baumgart C.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
DR   EMBL; AC123513; AAM44373.1; -; Genomic_DNA.
DR   EMBL; AC117070; AAM09345.2; -; Genomic_DNA.
DR   EMBL; AAFI01000027; EAL70168.1; -; Genomic_DNA.
DR   HSSP; P11376; 2EQL.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR001916; Glyco_hydro_22.
DR   InterPro; IPR008258; LT_catalytic.
DR   Pfam; PF01464; SLT; 1.
DR   SMART; SM00263; LYZ1; 1.
KW   Glycosidase.
SQ   SEQUENCE   202 AA;  19375 MW;  3905406541395584 CRC64;

Alignment Scores:
Pred. No.:              110           Length:         202
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x Q8MML4_DICDI (1-202)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        146 SerSerGlySerSerSerGlySerSer 154

RESULT 71
Q5AK97_CANAL
ID   Q5AK97_CANAL PRELIMINARY;      PRT;   208 AA.
AC   Q5AK97;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=CaO19.3923;
OS   Candida albicans SC5314.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=237561;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RX   PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA   Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA   Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA   Davis R.W., Scherer S.;
RT   "The diploid genome sequence of Candida albicans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
```

```
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RA   Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA   Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA   Jones T., Scherer S., Agabian N.;
RT   "Annotation of the Genome of Candida albicans.";
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACQ01000012; EAL03053.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   208 AA;  20615 MW;  1BFEA3E182CF609C CRC64;

Alignment Scores:
Pred. No.:              110           Length:        208
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5AK97_CANAL (1-208)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db         53 SerSerGlySerSerSerGlySerSer 61

RESULT 72
Q5AMW3_CANAL
ID   Q5AMW3_CANAL PRELIMINARY;      PRT;   208 AA.
AC   Q5AMW3;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=CaO19.12158;
OS   Candida albicans SC5314.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=237561;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RX   PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA   Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA   Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA   Davis R.W., Scherer S.;
RT   "The diploid genome sequence of Candida albicans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RA   Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA   Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA   Jones T., Scherer S., Agabian N.;
RT   "Annotation of the Genome of Candida albicans.";
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACQ01000005; EAL04178.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   208 AA;  20643 MW;  D90FAACEEF2D9F47 CRC64;

Alignment Scores:
Pred. No.:              110           Length:        208
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5AMW3_CANAL (1-208)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db         64 SerSerGlySerSerSerGlySerSer 72

RESULT 73
Q5P3Y9_AZOSE
ID   Q5P3Y9_AZOSE PRELIMINARY;      PRT;   208 AA.
AC   Q5P3Y9;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=AZOSEA18490; ORFNames=ebA3285;
OS   Azoarcus sp. (strain EbN1).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC   Rhodocyclaceae; Azoarcus.
OX   NCBI_TaxID=76114;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=EbN1;
RX   PubMed=15551059; DOI=10.1007/s00203-004-0742-9;
RA   Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,
RA   Reinhardt R.;
RT   "The genome sequence of an anaerobic aromatic-degrading denitrifying
RT   bacterium, strain EbN1.";
RL   Arch. Microbiol. 183:27-36(2005).
DR   EMBL; CR555306; CAI07974.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   208 AA;  22607 MW;  827F99AE49FB5AA6 CRC64;

Alignment Scores:
Pred. No.:              110           Length:        208
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5P3Y9_AZOSE (1-208)

Qy        229 TCCATTTCCAGTGCCAGAAGCTCTGGT 255
              ||||||||||||||||||||||||||
Db        174 SerIleSerSerAlaArgSerSerGly 182

RESULT 74
Q5AMF7_CANAL
ID   Q5AMF7_CANAL PRELIMINARY;      PRT;   214 AA.
AC   Q5AMF7;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=CaO19.4689;
OS   Candida albicans SC5314.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=237561;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RX   PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA   Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA   Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA   Davis R.W., Scherer S.;
RT   "The diploid genome sequence of Candida albicans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=SC5314;
RA   Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
```

```
RA   Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA   Jones T., Scherer S., Agabian N.;
RT   "Annotation of the Genome of Candida albicans.";
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACQ01000006; EAL04024.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   214 AA;  21206 MW;  8DC137D89BBB93EC CRC64;

Alignment Scores:
Pred. No.:              110           Length:        214
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q5AMF7_CANAL (1-214)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
                  ||||||||||||||||||||||||||
Db         75 SerSerGlySerSerSerGlySerSer 83

RESULT 75
Q82GT4_STRAW
ID   Q82GT4_STRAW PRELIMINARY;      PRT;   215 AA.
AC   Q82GT4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SAV3812;
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR   EMBL; BA000030; BAC71524.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   215 AA;  21931 MW;  EA74FA071FEBBB47 CRC64;

Alignment Scores:
Pred. No.:              110           Length:        215
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x Q82GT4_STRAW (1-215)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
                  ||||||||||||||||||||||||||
Db         91 SerGlySerSerSerSerGlySerSer 99

Search completed: March 21, 2006, 22:26:54
Job time : 248 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:14:09 ; Search time 29.3 Seconds
(without alignments)
4768.684 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4589790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152102_26777/app_query.fasta_1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=oligo.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=50 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_157_@runat_21032006_152102_26777 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	346	65.5	529	2	AAW80498	Aaw80498 A protein
2	346	65.5	529	2	AAW85055	Aaw85055 Human pro
3	346	65.5	529	4	AAU27550	Aau27550 Corneodes
4	346	65.5	529	7	ABJ39148	Abj39148 Human cor
5	346	65.5	529	8	ADN04226	Adn04226 Antipsori
6	245	46.4	529	8	ADP12525	Adp12525 Protein e
7	224	42.4	486	7	ABJ39147	Abj39147 Human S p
8	224	42.4	486	8	ADQ21504	Adq21504 Human sof
9	205	38.8	529	3	AAY44989	Aay44989 Human epi
10	123	23.3	529	4	AAG64152	Aag64152 Polypepti
11	94	17.8	219	8	ABO60412	Abo60412 Human gen
12	76	14.4	76	7	ABJ39197	Abj39197 Fragment
13	69	13.1	71	5	AAU91255	Aau91255 Nucleic a
14	69	13.1	146	7	ABJ39195	Abj39195 Fragment
15	59	11.2	111	7	ABJ39196	Abj39196 Fragment
16	58	11.0	58	7	ABJ39194	Abj39194 Fragment
17	55	10.4	55	7	ABJ39193	Abj39193 Fragment
18	22	4.2	35	5	ABG68620	Abg68620 Corneodes
19	21	4.0	21	5	ABG68622	Abg68622 Corneodes
20	15	2.8	15	7	ABJ39150	Abj39150 Serine ri
21	15	2.8	15	7	ABJ39152	Abj39152 Serine ri
22	14	2.7	14	5	ABG68623	Abg68623 Corneodes
23	14	2.7	16	7	ABJ39149	Abj39149 Signal pe
24	14	2.7	35	5	ABG68621	Abg68621 Corneodes
25	13	2.5	14	5	ABG68628	Abg68628 Human cor
26	11	2.1	11	7	ABJ39157	Abj39157 Serine ri
27	11	2.1	12	5	ABG68624	Abg68624 Corneodes
28	11	2.1	117	4	AAU32566	Aau32566 Novel hum
29	11	2.1	344	5	AAM51980	Aam51980 Maize bio
30	11	2.1	344	5	AAM51974	Aam51974 Maize bio
31	11	2.1	344	7	ADE76382	Ade76382 Corn SID2
32	11	2.1	344	7	ADE76370	Ade76370 Corn biot
33	11	2.1	374	3	AAG44803	Aag44803 Zea mays
34	11	2.1	377	3	AAG44805	Aag44805 Zea mays
35	11	2.1	377	5	AAM51982	Aam51982 Maize bio
36	11	2.1	377	5	AAM51981	Aam51981 Maize bio
37	11	2.1	377	7	ADE76384	Ade76384 Corn SID2
38	11	2.1	377	7	ADE76386	Ade76386 Corn SID2
39	11	2.1	381	8	ADY09732	Ady09732 Plant ful
40	11	2.1	631	6	AAO30442	Aao30442 Human MBM
41	11	2.1	631	7	ADJ70128	Adj70128 Human hea
42	11	2.1	651	8	ADO41710	Ado41710 Novel hum
43	11	2.1	709	6	ABJ19355	Abj19355 NOVX rela
44	11	2.1	709	8	ADQ67680	Adq67680 Novel hum
45	11	2.1	828	4	ABB68240	Abb68240 Drosophil
46	10	1.9	10	5	ABG68619	Abg68619 Human cor
47	10	1.9	309	8	ADN19766	Adn19766 Bacterial
48	10	1.9	344	8	ADN19767	Adn19767 Bacterial
49	10	1.9	348	2	AAR54665	Aar54665 Outer mem
50	10	1.9	406	7	ADB65137	Adb65137 Human pro

ALIGNMENTS

RESULT 1
AAW80498
ID AAW80498 standard; protein; 529 AA.
XX
AC AAW80498;
XX
DT 10-FEB-1999 (first entry)
XX
DE A protein designated corneodesmosin.
XX
KW Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW reactive keratosis; and leucokeratosis.
XX
OS Homo sapiens.
XX
PN FR2761362-A1.
XX
PD 02-OCT-1998.
XX
PF 28-MAR-1997; 97FR-00003899.
XX
PR 28-MAR-1997; 97FR-00003899.
XX
PA (OREA) L'OREAL SA.
XX
PI Serre G, Simon M, Weber VM;

```
XX
DR   WPI; 1998-586093/50.
DR   N-PSDB; AAV63614.
XX
PT   New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT   or therapeutic treatment of trophic disorders or thinning of the
PT   epidermis.
XX
PS   Claim 1; Page 29; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:                4.63e-302       Length:          529
Score:                    346.00          Matches:         346
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              65.5%           Indels:          0
DB:                       2               Gaps:            0

US-10-677-865A-2 (1-1590) x AAW80498 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 2
AAW85055
ID   AAW85055 standard; protein; 529 AA.
XX
AC   AAW85055;
XX
DT   05-FEB-1999  (first entry)
XX
DE   Human protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761363-A1.
XX
PD   02-OCT-1998.
XX
PF   11-SEP-1997;   97FR-00011317.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586094/50.
DR   N-PSDB; AAV63541.
XX
PT   New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT   cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT   epidermis etc., also compositions containing protease that degrades
PT   corneodesmosin.
XX
PS   Claim 1; Page 30; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
```

```
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              4.63e-302      Length:         529
Score:                  346.00         Matches:        346
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            65.5%          Indels:         0
DB:                     2              Gaps:           0

US-10-677-865A-2 (1-1590) x AAW85055 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 3
AAU27550
ID   AAU27550 standard; protein; 529 AA.
XX
AC   AAU27550;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
DR   N-PSDB; AAS43749.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Claim 8; Fig 2; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
```

```
CC   skin function. The present sequence represents corneodesmosin amino acid
CC   sequence, as described in the invention
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              4.63e-302      Length:         529
Score:                  346.00         Matches:        346
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            65.5%          Indels:         0
DB:                     4              Gaps:           0

US-10-677-865A-2 (1-1590) x AAU27550 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 4
ABJ39148
ID   ABJ39148 standard; protein; 529 AA.
XX
AC   ABJ39148;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human corneodesmosin (CDSN) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 28
FT                   /note= "Encoded by CCAG"
FT   misc_difference 408
FT                   /note= "Encoded by TCC"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   N-PSDB; ABT43849, ABT43850.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
```

```
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              4.63e-302       Length:         529
Score:                  346.00          Matches:        346
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            65.5%           Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x ABJ39148 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 5
ADN04226
ID   ADN04226 standard; protein; 529 AA.
XX
AC   ADN04226;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #308.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04225.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              4.63e-302       Length:         529
Score:                  346.00          Matches:        346
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            65.5%           Indels:         0
DB:                     8               Gaps:           0

US-10-677-865A-2 (1-1590) x ADN04226 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223
```

```
Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529
```

RESULT 6
ADP12525
ID ADP12525 standard; protein; 529 AA.
XX
AC ADP12525;
XX
DT 12-AUG-2004 (first entry)
XX
DE Protein encoded by mRNA of the invention #135.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.
XX
PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX
PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
DR WPI; 2004-400724/37.
XX
PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 529 AA;

Alignment Scores:

Pred. No.:	3.67e-211	Length:	529
Score:	245.00	Matches:	345
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	46.4%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADP12525 (1-529)

```
Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
```

```
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerSerGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529
```

RESULT 7
ABJ39147
ID ABJ39147 standard; protein; 486 AA.
XX
AC ABJ39147;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human S protein encoded by the skin corneodesmosin (CDSN) gene (S-gene).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW S protein; skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43848.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC polypeptide sequence is the human S protein (corneodesmosin protein only
CC expressed in the skin), the target of the invention
XX
SQ Sequence 486 AA;

Alignment Scores:

Pred. No.:	2.95e-192	Length:	486
Score:	224.00	Matches:	224
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	42.4%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39147 (1-486)

```
Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 187

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 207

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 227

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 247

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 267

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 287
```

```
Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 307

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 327

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 347

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 367

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 387

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        388 IleSerSerSer 391
```

RESULT 8
ADQ21504
ID ADQ21504 standard; protein; 486 AA.
XX
AC ADQ21504;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4324.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4324; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX

SQ Sequence 486 AA;

Alignment Scores:

Pred. No.:	2.95e-192	Length:	486
Score:	224.00	Matches:	224
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	42.4%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADQ21504 (1-486)

```
Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerGlnThrSerGly 187

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 207

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 227

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 247

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 267

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 287

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 307

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 327

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 347

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 367

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 387

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        388 IleSerSerSer 391
```

RESULT 9
AAY44989
ID AAY44989 standard; protein; 529 AA.
XX
AC AAY44989;
XX
DT 23-MAY-2000 (first entry)
XX
DE Human epidermal protein-6.
XX
KW Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; autoimmune disorder; inflammatory disorder;

```
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .32
FT                   /label= Signal_peptide
FT   Modified-site   3
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   37
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   50
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   57
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   80
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   172
FT                   /note= "N-glycosylated"
FT   Modified-site   182
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   287
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   395
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   440
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   N-PSDB; AAZ50582.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 1; Fig 4; 82pp; English.
XX
CC   The present sequence is human epidermal protein-6 (HEPI) expressed in
CC   breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC   cDNA library. HEPI-6 shares 99% identity with human S protein.
CC   Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC   for protein expression. The HEPI proteins are useful for the treatment of
CC   epithelial disorders, including dyshidrotic eczema and scabies, cell
CC   proliferative disorders including actinic keratosis and arteriosclerosis,
CC   and autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression
XX
SQ   Sequence 529 AA;
```

```
Alignment Scores:
Pred. No.:              3.67e-175      Length:          529
Score:                  205.00         Matches:         205
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            38.8%          Indels:          0
DB:                     3              Gaps:            0

US-10-677-865A-2 (1-1590) x AAY44989 (1-529)

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIle 222

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSer 242

Qy        727 TCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 SerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySer 262

Qy        787 GGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysPro 282

Qy        847 TGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySer 302

Qy        907 TCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGC 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 SerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGly 322

Qy        967 TACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGly 342

Qy       1027 CCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCA 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 ProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAla 362

Qy       1087 GCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGly 382

Qy       1147 GGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSer 402

Qy       1207 AGCATTTCCAGCAGC 1221
              |||||||||||||||
Db        403 SerIleSerSerSer 407

RESULT 10
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 143
FT                   /label= Ser, Asn
FT   Misc-difference 151
```

```
FT                   /label= Ser, Cys
FT   Misc-difference 152
FT                   /label= Ser, Arg
FT   Misc-difference 200
FT                   /label= Ser, Phe
FT   Misc-difference 282
FT                   /label= Pro, Thr
FT   Misc-difference 401
FT                   /label= Ser, Gly
FT   Misc-difference 408
FT                   /label= Ala, Ser
FT   Misc-difference 409
FT                   /label= Gly, Val
FT   Misc-difference 410
FT                   /label= Ser, Leu
FT   Misc-difference 453
FT                   /label= Ser, Asn
FT   Misc-difference 527
FT                   /label= Asp, Asn
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

Alignment Scores:
Pred. No.:              2.31e-101      Length:         529
Score:                  123.00         Matches:        123
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            23.3%          Indels:         0
DB:                     4              Gaps:           0

US-10-677-865A-2 (1-1590) x AAG64152 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCT 369
              |||||||||
Db        121 SerGlySer 123

RESULT 11
ABO60412
ID   ABO60412 standard; protein; 219 AA.
XX
AC   ABO60412;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon protein #6646.
XX
KW   Human; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 45; SEQ ID NO 34046; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
```

```
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe protein of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 219 AA;

Alignment Scores:
Pred. No.:              3.25e-75      Length:          219
Score:                  94.00         Matches:         94
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            17.8%         Indels:          0
DB:                     8             Gaps:            0

US-10-677-865A-2 (1-1590) x ABO60412 (1-219)

Qy         88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ThrLeuAlaLysSerIleGlyThrPheSerAspProCysLysAspProThrArgIleThr 20

Qy        148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSer 40

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlySerSerSerSerGlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSer 60

Qy        268 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGly 80

Qy        328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCT 369
              |||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySer 94

RESULT 12
ABJ39197
ID   ABJ39197 standard; protein; 76 AA.
XX
AC   ABJ39197;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment e, glycine rich sticky domain 2 of the human S protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 76 AA;

Alignment Scores:
Pred. No.:              5.88e-59      Length:          76
Score:                  76.00         Matches:         76
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            14.4%         Indels:          0
DB:                     7             Gaps:            0

US-10-677-865A-2 (1-1590) x ABJ39197 (1-76)

Qy       1123 GGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCC 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyGlyValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSer 20

Qy       1183 AGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCC 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerSerArgValProSerSerSerSerIleSerSerSerAlaGlySerProTyrHisPro 40

Qy       1243 TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 CysGlySerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSer 60

Qy       1303 TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGC 1350
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerSerSerGlnSerSerGlyLysIleIleLeuGlnProCysGlySer 76

RESULT 13
AAU91255
ID   AAU91255 standard; protein; 71 AA.
XX
AC   AAU91255;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Nucleic acid selection related corneodesmosin peptide.
XX
KW   Nucleic acid selection; viral genome; sindbis virus.
XX
OS   Mus musculus.
XX
PN   WO200216572-A2.
```

XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-CA001169.
XX
PR 18-AUG-2000; 2000US-00641931.
XX
PA (PHEN-) PHENOGENE THERAPEUTIQUES INC.
XX
PI Lanctot C, Moffat P, Salois P;
XX
DR WPI; 2002-257913/30.
XX
PT Selecting exogenous nucleic acids having a desired feature, particularly
PT from a library of nucleic acids, involves using a viral genome.
XX
PS Example 4; Page 55; 88pp; English.
XX
CC The invention describes the use of a viral genome for selecting an
CC exogenous nucleic acid having a desired feature. The method requires a
CC viral genome, e.g. sindbis virus, capable of expressing the exogenous
CC nucleic acid in a host cell and selecting among these the ones having the
CC desired feature or predefined function. The method removes the expensive
CC and time-consuming task of selecting cells that express a gene of
CC interest. The method is more rapid, efficient and accurate for selecting
CC a particular nucleic acid having a desired feature, characteristic or
CC function. This sequence represents a peptide encoded by a nucleic acid
CC isolated using the method described in the invention
XX
SQ Sequence 71 AA;

Alignment Scores:

Pred. No.:	1.18e-52	Length:	71
Score:	69.00	Matches:	69
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	13.1%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAU91255 (1-71)

```
Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGT 207
              |||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSer 69
```

RESULT 14
ABJ39195
ID ABJ39195 standard; protein; 146 AA.
XX
AC ABJ39195;
XX
DT 06-NOV-2003 (first entry)
XX
DE Fragment c, cleavage product of human S (corneodesmosin skin) protein.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 146 AA;

Alignment Scores:

Pred. No.:	1.08e-52	Length:	146
Score:	69.00	Matches:	69
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	13.1%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39195 (1-146)

```
Qy        163 TGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCT 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSer 20

Qy        223 GGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGA 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGly 40

Qy        283 TCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAG 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGln 60

Qy        343 GTCAGCTACTCCTCCGGATCTGGCTCT 369
              |||||||||||||||||||||||||||
Db         61 ValSerTyrSerSerGlySerGlySer 69
```

RESULT 15
ABJ39196
ID ABJ39196 standard; protein; 111 AA.
XX
AC ABJ39196;
XX

DT 06-NOV-2003 (first entry)
XX
DE Fragment d, glycine rich sticky domain 1 of the human S protein.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 111 AA;

Alignment Scores:

Pred. No.:	1.12e-43	Length:	111
Score:	59.00	Matches:	59
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.2%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39196 (1-111)

```
Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSerAlaArgSerSer 20

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGly 40

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCT 369
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySer 59
```

RESULT 16
ABJ39194
ID ABJ39194 standard; protein; 58 AA.
XX
AC ABJ39194;
XX
DT 06-NOV-2003 (first entry)
XX
DE Fragment b, cleavage product of human S (corneodesmosin skin) protein.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 58 AA;

Alignment Scores:

Pred. No.:	9.64e-43	Length:	58
Score:	58.00	Matches:	58
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.0%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39194 (1-58)

```
Qy       1414 GATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspGlySerProHisProAspProSerAlaGlyAlaLysProCysGlySerSerSerAla 20

Qy       1474 GGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         21 GlyLysIleProCysArgSerIleArgAspIleLeuAlaGlnValLysProLeuGlyPro 40

Qy       1534 CAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlnLeuAlaAspProGluValPheLeuProGlnGlyGluLeuLeuAspSerPro 58

RESULT 17
ABJ39193
ID   ABJ39193 standard; protein; 55 AA.
XX
AC   ABJ39193;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Fragment a, cleavage product of human S (corneodesmosin skin) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 55 AA;

Alignment Scores:
Pred. No.:              4.86e-40        Length:         55
Score:                  55.00           Matches:        55
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            10.4%           Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x ABJ39193 (1-55)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGC 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCys 55

RESULT 18
ABG68620
ID   ABG68620 standard; peptide; 35 AA.
XX
AC   ABG68620;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Corneodesmosin peptide #1.
XX
KW   Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW   stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW   desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW   contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW   bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW   Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW   impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW   basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW   gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS   Unidentified.
XX
PN   WO200244736-A2.
XX
PD   06-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-GB005303.
XX
PR   30-NOV-2000; 2000GB-00029225.
PR   07-DEC-2000; 2000GB-00029879.
XX
PA   (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI   Tazi-Ahnini R,  Bavik C,  Ward S,  Duff G,  Cork M;
XX
DR   WPI; 2002-557554/59.
XX
PT   Diagnosing a disease or susceptibility to a disease associated with
PT   abnormal cell-cell adhesion between epithelial cells, by detecting
PT   mutation in nucleic acid encoding adhesion protein, protease or protease
PT   inhibitor.
XX
PS   Example A5; Page 102; 257pp; English.
XX
CC   The invention relates to a method for diagnosis of a disease or
CC   susceptibility to a disease associated with abnormal cell-cell adhesion
CC   between epithelial cells, comprising detecting a mutation in a nucleic
CC   acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC   modulated level of adhesion protein, protease or protease inhibitor, or
CC   its fragment polypeptide. The method is useful for diagnosing a disease
CC   such as eczema, contact dermatitis, lung atopic asthma, post viral
CC   asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC   disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC   ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC   cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC   malignancy of the gastrointestinal tract, malignancy of the lung,
CC   psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC   a peptide used in the scope of the invention
XX
SQ   Sequence 35 AA;
```

```
Alignment Scores:
Pred. No.:              2.58e-10        Length:         35
Score:                  22.00           Matches:        22
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            4.2%            Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x ABG68620 (1-35)

Qy        133 CCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGC 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGly 20

Qy        193 TTCAGT 198
              ||||||
Db         21 PheSer 22

RESULT 19
ABG68622
ID   ABG68622 standard; peptide; 21 AA.
XX
AC   ABG68622;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Corneodesmosin peptide #3.
XX
KW   Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW   stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW   desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW   contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW   bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW   Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW   impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW   basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW   gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS   Unidentified.
XX
PN   WO200244736-A2.
XX
PD   06-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-GB005303.
XX
PR   30-NOV-2000; 2000GB-00029225.
PR   07-DEC-2000; 2000GB-00029879.
XX
PA   (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI   Tazi-Ahnini R,  Bavik C,  Ward S,  Duff G,  Cork M;
XX
DR   WPI; 2002-557554/59.
XX
PT   Diagnosing a disease or susceptibility to a disease associated with
PT   abnormal cell-cell adhesion between epithelial cells, by detecting
PT   mutation in nucleic acid encoding adhesion protein, protease or protease
PT   inhibitor.
XX
PS   Example A5; Page 102; 257pp; English.
XX
CC   The invention relates to a method for diagnosis of a disease or
CC   susceptibility to a disease associated with abnormal cell-cell adhesion
CC   between epithelial cells, comprising detecting a mutation in a nucleic
CC   acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC   modulated level of adhesion protein, protease or protease inhibitor, or
CC   its fragment polypeptide. The method is useful for diagnosing a disease
CC   such as eczema, contact dermatitis, lung atopic asthma, post viral
CC   asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC   disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC   ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC   cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC   malignancy of the gastrointestinal tract, malignancy of the lung,
CC   psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC   a peptide used in the scope of the invention
XX
SQ   Sequence 21 AA;

Alignment Scores:
Pred. No.:              2.18e-09        Length:         21
Score:                  21.00           Matches:        21
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            4.0%            Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x ABG68622 (1-21)

Qy        133 CCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGC 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGly 20

Qy        193 TTC 195
              |||
Db         21 Phe 21

RESULT 20
ABJ39150
ID   ABJ39150 standard; peptide; 15 AA.
XX
AC   ABJ39150;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Serine rich domain 1 of the human S protein (skin corneodesmosin).
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2b; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
```

CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 15 AA;

Alignment Scores:

Pred. No.:	0.000572	Length:	15
Score:	15.00	Matches:	15
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.8%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39150 (1-15)

```
Qy        196 AGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 15
```

RESULT 21
ABJ39152
ID ABJ39152 standard; peptide; 15 AA.
XX
AC ABJ39152;
XX
DT 06-NOV-2003 (first entry)
XX
DE Serine rich domain 3 of the human S protein (skin corneodesmosin).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 15 AA;

Alignment Scores:

Pred. No.:	0.000572	Length:	15
Score:	15.00	Matches:	15
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.8%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39152 (1-15)

```
Qy        436 TCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 15
```

RESULT 22
ABG68623
ID ABG68623 standard; peptide; 14 AA.
XX
AC ABG68623;
XX
DT 07-OCT-2002 (first entry)
XX
DE Corneodesmosin peptide #4.
XX
KW Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS Unidentified.
XX
PN WO200244736-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-GB005303.
XX
PR 30-NOV-2000; 2000GB-00029225.
PR 07-DEC-2000; 2000GB-00029879.
XX
PA (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI Tazi-Ahnini R, Bavik C, Ward S, Duff G, Cork M;
XX
DR WPI; 2002-557554/59.
XX
PT Diagnosing a disease or susceptibility to a disease associated with
PT abnormal cell-cell adhesion between epithelial cells, by detecting
PT mutation in nucleic acid encoding adhesion protein, protease or protease
PT inhibitor.
XX
PS Example A5; Page 102; 257pp; English.
XX
CC The invention relates to a method for diagnosis of a disease or
CC susceptibility to a disease associated with abnormal cell-cell adhesion
CC between epithelial cells, comprising detecting a mutation in a nucleic
CC acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC modulated level of adhesion protein, protease or protease inhibitor, or
CC its fragment polypeptide. The method is useful for diagnosing a disease

CC such as eczema, contact dermatitis, lung atopic asthma, post viral
CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC malignancy of the gastrointestinal tract, malignancy of the lung,
CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC a peptide used in the scope of the invention
XX
SQ Sequence 14 AA;

Alignment Scores:

Pred. No.:	0.00459	Length:	14
Score:	14.00	Matches:	14
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.7%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x ABG68623 (1-14)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGA 246
              ||||||||||||||||||||||||||||||||||||||||||
Db          1 SerGlySerSerSerSerGlySerSerIleSerSerAlaArg 14
```

RESULT 23
ABJ39149
ID ABJ39149 standard; peptide; 16 AA.
XX
AC ABJ39149;
XX
DT 06-NOV-2003 (first entry)
XX
DE Signal peptide of the human S protein (skin corneodesmosin).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 16 AA;

Alignment Scores:

Pred. No.:	0.00451	Length:	16
Score:	14.00	Matches:	14
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39149 (1-16)

```
Qy         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCT 96
              ||||||||||||||||||||||||||||||||||||||||||
Db          3 AlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAla 16
```

RESULT 24
ABG68621
ID ABG68621 standard; peptide; 35 AA.
XX
AC ABG68621;
XX
DT 07-OCT-2002 (first entry)
XX
DE Corneodesmosin peptide #2.
XX
KW Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS Unidentified.
XX
PN WO200244736-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-GB005303.
XX
PR 30-NOV-2000; 2000GB-00029225.
PR 07-DEC-2000; 2000GB-00029879.
XX
PA (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI Tazi-Ahnini R, Bavik C, Ward S, Duff G, Cork M;
XX
DR WPI; 2002-557554/59.
XX
PT Diagnosing a disease or susceptibility to a disease associated with
PT abnormal cell-cell adhesion between epithelial cells, by detecting
PT mutation in nucleic acid encoding adhesion protein, protease or protease
PT inhibitor.
XX
PS Example A5; Page 102; 257pp; English.
XX
CC The invention relates to a method for diagnosis of a disease or
CC susceptibility to a disease associated with abnormal cell-cell adhesion
CC between epithelial cells, comprising detecting a mutation in a nucleic
CC acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC modulated level of adhesion protein, protease or protease inhibitor, or

```
CC   its fragment polypeptide. The method is useful for diagnosing a disease
CC   such as eczema, contact dermatitis, lung atopic asthma, post viral
CC   asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC   disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC   ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC   cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC   malignancy of the gastrointestinal tract, malignancy of the lung,
CC   psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC   a peptide used in the scope of the invention
XX
SQ   Sequence 35 AA;

Alignment Scores:
Pred. No.:              0.00409        Length:         35
Score:                  14.00          Matches:        14
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            2.7%           Indels:         0
DB:                     5              Gaps:           0

US-10-677-865A-2 (1-1590) x ABG68621 (1-35)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGA 246
              ||||||||||||||||||||||||||||||||||||||||||
Db         22 SerGlySerSerSerSerGlySerSerIleSerSerAlaArg 35

RESULT 25
ABG68628
ID   ABG68628 standard; peptide; 14 AA.
XX
AC   ABG68628;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human corneodesmosin peptide #2.
XX
KW   Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW   stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW   desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW   contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW   bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW   Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW   impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW   basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW   gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS   Homo sapiens.
XX
PN   WO200244736-A2.
XX
PD   06-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-GB005303.
XX
PR   30-NOV-2000; 2000GB-00029225.
PR   07-DEC-2000; 2000GB-00029879.
XX
PA   (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI   Tazi-Ahnini R,  Bavik C,  Ward S,  Duff G,  Cork M;
XX
DR   WPI; 2002-557554/59.
XX
PT   Diagnosing a disease or susceptibility to a disease associated with
PT   abnormal cell-cell adhesion between epithelial cells, by detecting
PT   mutation in nucleic acid encoding adhesion protein, protease or protease
PT   inhibitor.
XX
PS   Example C4; Page 145; 257pp; English.
XX
CC   The invention relates to a method for diagnosis of a disease or
CC   susceptibility to a disease associated with abnormal cell-cell adhesion
CC   between epithelial cells, comprising detecting a mutation in a nucleic
CC   acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC   modulated level of adhesion protein, protease or protease inhibitor, or
CC   its fragment polypeptide. The method is useful for diagnosing a disease
CC   such as eczema, contact dermatitis, lung atopic asthma, post viral
CC   asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC   disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC   ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC   cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC   malignancy of the gastrointestinal tract, malignancy of the lung,
CC   psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC   a peptide used in the scope of the invention
XX
SQ   Sequence 14 AA;

Alignment Scores:
Pred. No.:              0.0364         Length:         14
Score:                  13.00          Matches:        13
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            2.5%           Indels:         0
DB:                     5              Gaps:           0

US-10-677-865A-2 (1-1590) x ABG68628 (1-14)

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTC 1008
              |||||||||||||||||||||||||||||||||||||||
Db          1 PheThrLysGluAsnProValLysGlySerProGlyVal 13

RESULT 26
ABJ39157
ID   ABJ39157 standard; peptide; 11 AA.
XX
AC   ABJ39157;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Serine rich domain 7 of the human S protein (skin corneodesmosin).
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW   skin.
XX
OS   Homo sapiens.
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2b; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
```

CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 11 AA;

Alignment Scores:

Pred. No.:	2.37	Length:	11
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39157 (1-11)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||||||||
Db          1 SerPheSerSerSerSerSerSerGlnSerSer 11
```

RESULT 27
ABG68624
ID ABG68624 standard; peptide; 12 AA.
XX
AC ABG68624;
XX
DT 07-OCT-2002 (first entry)
XX
DE Corneodesmosin peptide #5.
XX
KW Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS Unidentified.
XX
PN WO200244736-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-GB005303.
XX
PR 30-NOV-2000; 2000GB-00029225.
PR 07-DEC-2000; 2000GB-00029879.
XX
PA (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI Tazi-Ahnini R, Bavik C, Ward S, Duff G, Cork M;
XX
DR WPI; 2002-557554/59.
XX
PT Diagnosing a disease or susceptibility to a disease associated with
PT abnormal cell-cell adhesion between epithelial cells, by detecting
PT mutation in nucleic acid encoding adhesion protein, protease or protease
PT inhibitor.
XX
PS Example A5; Page 102; 257pp; English.
XX
CC The invention relates to a method for diagnosis of a disease or
CC susceptibility to a disease associated with abnormal cell-cell adhesion
CC between epithelial cells, comprising detecting a mutation in a nucleic
CC acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC modulated level of adhesion protein, protease or protease inhibitor, or
CC its fragment polypeptide. The method is useful for diagnosing a disease
CC such as eczema, contact dermatitis, lung atopic asthma, post viral
CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC malignancy of the gastrointestinal tract, malignancy of the lung,
CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC a peptide used in the scope of the invention
XX
SQ Sequence 12 AA;

Alignment Scores:

Pred. No.:	2.34	Length:	12
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x ABG68624 (1-12)

```
Qy        133 CCCACGCGTATCACCTCCCCTAACGACCCCTGC 165
              ||||||||||||||||||||||||||||||||||
Db          1 ProThrArgIleThrSerProAsnAspProCys 11
```

RESULT 28
AAU32566
ID AAU32566 standard; protein; 117 AA.
XX
AC AAU32566;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3057.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 632; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 117 AA;

Alignment Scores:

Pred. No.:	1.76	Length:	117
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x AAU32566 (1-117)

```
Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||
Db         61 SerSerSerSerHisSerSerSerSerSerSer 71
```

RESULT 29
AAM51980
ID AAM51980 standard; protein; 344 AA.
XX
AC AAM51980;
XX
DT 06-FEB-2002 (first entry)
XX
DE Maize biotin synthase protein #4.
XX
KW Maize; biotin synthase; biotin biosynthesis; herbicide; tassel;
KW plant development; clone cdt2c.pk002.c17; transgenic plant.
XX
OS Zea mays.
XX
PN US2001039042-A1.
XX
PD 08-NOV-2001.
XX
PF 19-DEC-2000; 2000US-00740288.
XX
PR 21-DEC-1999; 99US-0172929P.
XX
PA (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO E M.
XX
PI Allen SM, Kinney AJ, Miao G, Orozco EM;
XX
DR WPI; 2002-040723/05.
DR N-PSDB; ABA01973.
XX
PT New polypeptides, useful as targets for herbicide discovery, and as
PT probes for genetic and physical mapping of genes of which they are part,
PT or creating transgenic plants, comprises biotin synthase polypeptides and
PT encoding polynucleotides.
XX
PS Claim 19; Fig 1; 46pp; English.
XX
CC The present invention provides the protein and coding sequences of biotin
CC synthase enzymes from barley, wheat, maize, soybean and the prickly
CC poppy. These sequences can be used to produce transgenic plants which
CC express different levels of the gene, or express it at different times in
CC plant development. They can also be used as a target in the production of
CC herbicides. The present sequence is a maize biotin synthase protein
CC obtained from clone cdt2c.pk002.c17, which was derived from a maize
CC developing tassel
XX
SQ Sequence 344 AA;

Alignment Scores:

Pred. No.:	1.54	Length:	344
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAM51980 (1-344)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 30
AAM51974
ID AAM51974 standard; protein; 344 AA.
XX
AC AAM51974;
XX
DT 06-FEB-2002 (first entry)
XX
DE Maize biotin synthase protein #3.
XX
KW Maize; biotin synthase; biotin biosynthesis; herbicide;
KW plant development; transgenic plant.
XX
OS Zea mays.
XX
PN US2001039042-A1.
XX
PD 08-NOV-2001.
XX
PF 19-DEC-2000; 2000US-00740288.
XX
PR 21-DEC-1999; 99US-0172929P.
XX
PA (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO E M.
XX
PI Allen SM, Kinney AJ, Miao G, Orozco EM;
XX
DR WPI; 2002-040723/05.
DR N-PSDB; ABA01967.
XX
PT New polypeptides, useful as targets for herbicide discovery, and as
PT probes for genetic and physical mapping of genes of which they are part,
PT or creating transgenic plants, comprises biotin synthase polypeptides and
PT encoding polynucleotides.
XX
PS Claim 19; Page 18-19; 46pp; English.
XX
CC The present invention provides the protein and coding sequences of biotin
CC synthase enzymes from barley, wheat, maize, soybean and the prickly
CC poppy. These sequences can be used to produce transgenic plants which
CC express different levels of the gene, or express it at different times in
CC plant development. They can also be used as a target in the production of
CC herbicides. The present sequence is a maize biotin synthase protein
XX
SQ Sequence 344 AA;

Alignment Scores:

Pred. No.:	1.54	Length:	344
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0

```
Query Match:            2.1%          Indels:        0
DB:                     5             Gaps:          0

US-10-677-865A-2 (1-1590) x AAM51974 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 31
ADE76382
ID   ADE76382 standard; protein; 344 AA.
XX
AC   ADE76382;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Corn SID20 biotin synthase protein - SEQ ID 20.
XX
KW   biotin synthase; plant; breeding; seed; herbicide; enzyme; maize; corn;
KW   SID20.
XX
OS   Zea mays.
XX
PN   US2003192073-A1.
XX
PD   09-OCT-2003.
XX
PF   07-MAY-2003; 2003US-00431544.
XX
PR   21-DEC-1999;   99US-0172929P.
PR   19-DEC-2000; 2000US-00740288.
XX
PA   (ALLE/) ALLEN S M.
PA   (KINN/) KINNEY A J.
PA   (MIAO/) MIAO G.
PA   (OROZ/) OROZCO E M.
XX
PI   Allen SM,  Kinney AJ,  Miao G,  Orozco EM;
XX
DR   WPI; 2003-831831/77.
DR   N-PSDB; ADE76381.
XX
PT   New isolated polynucleotide encoding plant biotin synthase, useful in
PT   plant molecular biology and plant breeding purposes, particularly in the
PT   production of altered levels of biotin synthase in plants and seeds.
XX
PS   Claim 19; SEQ ID NO 20; 47pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide encoding a first
CC   polypeptide or a third polypeptide where the first or third polypeptide
CC   is a biotin synthase. The methods and compositions of the invention may
CC   be useful in plant molecular biology and breeding, particularly in the
CC   production of altered levels of biotin synthase in plants and seeds. The
CC   polypeptides may be used as a target to facilitate the design and/or
CC   identification of inhibitors of these enzymes that may be useful as
CC   herbicides. Furthermore, the polynucleotides may be used as probes for
CC   genetically and physically mapping the genes of which they are a part.
CC   The current sequence is that of the corn biotin synthase-related protein
CC   of the invention.
XX
SQ   Sequence 344 AA;

Alignment Scores:
Pred. No.:              1.54          Length:        344
Score:                  11.00         Matches:       11
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            2.1%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x ADE76382 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 32
ADE76370
ID   ADE76370 standard; protein; 344 AA.
XX
AC   ADE76370;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Corn biotin synthase-related protein - SEQ ID 8.
XX
KW   biotin synthase; plant; breeding; seed; herbicide; enzyme; maize; corn.
XX
OS   Zea mays.
XX
PN   US2003192073-A1.
XX
PD   09-OCT-2003.
XX
PF   07-MAY-2003; 2003US-00431544.
XX
PR   21-DEC-1999;   99US-0172929P.
PR   19-DEC-2000; 2000US-00740288.
XX
PA   (ALLE/) ALLEN S M.
PA   (KINN/) KINNEY A J.
PA   (MIAO/) MIAO G.
PA   (OROZ/) OROZCO E M.
XX
PI   Allen SM,  Kinney AJ,  Miao G,  Orozco EM;
XX
DR   WPI; 2003-831831/77.
DR   N-PSDB; ADE76369.
XX
PT   New isolated polynucleotide encoding plant biotin synthase, useful in
PT   plant molecular biology and plant breeding purposes, particularly in the
PT   production of altered levels of biotin synthase in plants and seeds.
XX
PS   Claim 19; SEQ ID NO 8; 47pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide encoding a first
CC   polypeptide or a third polypeptide where the first or third polypeptide
CC   is a biotin synthase. The methods and compositions of the invention may
CC   be useful in plant molecular biology and breeding, particularly in the
CC   production of altered levels of biotin synthase in plants and seeds. The
CC   polypeptides may be used as a target to facilitate the design and/or
CC   identification of inhibitors of these enzymes that may be useful as
CC   herbicides. Furthermore, the polynucleotides may be used as probes for
CC   genetically and physically mapping the genes of which they are a part.
CC   The current sequence is that of the corn biotin synthase-related protein
CC   of the invention.
XX
SQ   Sequence 344 AA;

Alignment Scores:
Pred. No.:              1.54          Length:        344
Score:                  11.00         Matches:       11
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            2.1%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x ADE76370 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 33
AAG44803
```

ID AAG44803 standard; protein; 374 AA.
XX
AC AAG44803;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 56165.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	1.52	Length:	374
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x AAG44803 (1-374)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         16 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 26
```

RESULT 34
AAG44805
ID AAG44805 standard; protein; 377 AA.
XX
AC AAG44805;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 56167.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	1.52	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x AAG44805 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              ||||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 35
AAM51982
ID AAM51982 standard; protein; 377 AA.
XX
AC AAM51982;
XX

DT 06-FEB-2002 (first entry)
XX
DE Maize biotin synthase protein #6.
XX
KW Maize; biotin synthase; biotin biosynthesis; herbicide;
KW plant development; clone cca.pk0012.g11:fis; transgenic plant.
XX
OS Zea mays.
XX
PN US2001039042-A1.
XX
PD 08-NOV-2001.
XX
PF 19-DEC-2000; 2000US-00740288.
XX
PR 21-DEC-1999; 99US-0172929P.
XX
PA (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO E M.
XX
PI Allen SM, Kinney AJ, Miao G, Orozco EM;
XX
DR WPI; 2002-040723/05.
DR N-PSDB; ABA01975.
XX
PT New polypeptides, useful as targets for herbicide discovery, and as
PT probes for genetic and physical mapping of genes of which they are part,
PT or creating transgenic plants, comprises biotin synthase polypeptides and
PT encoding polynucleotides.
XX
PS Claim 19; Fig 1; 46pp; English.
XX
CC The present invention provides the protein and coding sequences of biotin
CC synthase enzymes from barley, wheat, maize, soybean and the prickly
CC poppy. These sequences can be used to produce transgenic plants which
CC express different levels of the gene, or express it at different times in
CC plant development. They can also be used as a target in the production of
CC herbicides. The present sequence is a maize biotin synthase protein
CC obtained from clone cca.pk0012.g11:fis
XX
SQ Sequence 377 AA;

Alignment Scores:

Pred. No.:	1.52	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAM51982 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 36
AAM51981
ID AAM51981 standard; protein; 377 AA.
XX
AC AAM51981;
XX
DT 06-FEB-2002 (first entry)
XX
DE Maize biotin synthase protein #5.
XX
KW Maize; biotin synthase; biotin biosynthesis; herbicide;
KW plant development; clone cho1c.pk009.j14; transgenic plant.
XX
OS Zea mays.
XX
PN US2001039042-A1.
XX
PD 08-NOV-2001.
XX
PF 19-DEC-2000; 2000US-00740288.
XX
PR 21-DEC-1999; 99US-0172929P.
XX
PA (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO E M.
XX
PI Allen SM, Kinney AJ, Miao G, Orozco EM;
XX
DR WPI; 2002-040723/05.
DR N-PSDB; ABA01974.
XX
PT New polypeptides, useful as targets for herbicide discovery, and as
PT probes for genetic and physical mapping of genes of which they are part,
PT or creating transgenic plants, comprises biotin synthase polypeptides and
PT encoding polynucleotides.
XX
PS Claim 19; Fig 1; 46pp; English.
XX
CC The present invention provides the protein and coding sequences of biotin
CC synthase enzymes from barley, wheat, maize, soybean and the prickly
CC poppy. These sequences can be used to produce transgenic plants which
CC express different levels of the gene, or express it at different times in
CC plant development. They can also be used as a target in the production of
CC herbicides. The present sequence is a maize biotin synthase protein
CC obtained from clone cho1c.pk009.j14, which was derived from a maize
CC embryo
XX
SQ Sequence 377 AA;

Alignment Scores:

Pred. No.:	1.52	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAM51981 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 37
ADE76384
ID ADE76384 standard; protein; 377 AA.
XX
AC ADE76384;
XX
DT 29-JAN-2004 (first entry)
XX
DE Corn SID22 biotin synthase protein - SEQ ID 22.
XX
KW biotin synthase; plant; breeding; seed; herbicide; enzyme; maize; corn;
KW SID22.
XX
OS Zea mays.
XX
PN US2003192073-A1.
XX
PD 09-OCT-2003.
XX
PF 07-MAY-2003; 2003US-00431544.
XX
PR 21-DEC-1999; 99US-0172929P.
PR 19-DEC-2000; 2000US-00740288.

```
XX
PA   (ALLE/) ALLEN S M.
PA   (KINN/) KINNEY A J.
PA   (MIAO/) MIAO G.
PA   (OROZ/) OROZCO E M.
XX
PI   Allen SM,  Kinney AJ,  Miao G,  Orozco EM;
XX
DR   WPI; 2003-831831/77.
DR   N-PSDB; ADE76383.
XX
PT   New isolated polynucleotide encoding plant biotin synthase, useful in
PT   plant molecular biology and plant breeding purposes, particularly in the
PT   production of altered levels of biotin synthase in plants and seeds.
XX
PS   Claim 19; SEQ ID NO 22; 47pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide encoding a first
CC   polypeptide or a third polypeptide where the first or third polypeptide
CC   is a biotin synthase. The methods and compositions of the invention may
CC   be useful in plant molecular biology and breeding, particularly in the
CC   production of altered levels of biotin synthase in plants and seeds. The
CC   polypeptides may be used as a target to facilitate the design and/or
CC   identification of inhibitors of these enzymes that may be useful as
CC   herbicides. Furthermore, the polynucleotides may be used as probes for
CC   genetically and physically mapping the genes of which they are a part.
CC   The current sequence is that of the corn biotin synthase-related protein
CC   of the invention.
XX
SQ   Sequence 377 AA;

Alignment Scores:
Pred. No.:              1.52          Length:         377
Score:                  11.00         Matches:        11
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            2.1%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x ADE76384 (1-377)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 38
ADE76386
ID   ADE76386 standard; protein; 377 AA.
XX
AC   ADE76386;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Corn SID24 biotin synthase protein - SEQ ID 24.
XX
KW   biotin synthase; plant; breeding; seed; herbicide; enzyme; maize; corn;
KW   SID24.
XX
OS   Zea mays.
XX
PN   US2003192073-A1.
XX
PD   09-OCT-2003.
XX
PF   07-MAY-2003; 2003US-00431544.
XX
PR   21-DEC-1999;   99US-0172929P.
PR   19-DEC-2000; 2000US-00740288.
XX
PA   (ALLE/) ALLEN S M.
PA   (KINN/) KINNEY A J.
PA   (MIAO/) MIAO G.
PA   (OROZ/) OROZCO E M.
XX
PI   Allen SM,  Kinney AJ,  Miao G,  Orozco EM;
XX
DR   WPI; 2003-831831/77.
DR   N-PSDB; ADE76385.
XX
PT   New isolated polynucleotide encoding plant biotin synthase, useful in
PT   plant molecular biology and plant breeding purposes, particularly in the
PT   production of altered levels of biotin synthase in plants and seeds.
XX
PS   Claim 19; SEQ ID NO 24; 47pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide encoding a first
CC   polypeptide or a third polypeptide where the first or third polypeptide
CC   is a biotin synthase. The methods and compositions of the invention may
CC   be useful in plant molecular biology and breeding, particularly in the
CC   production of altered levels of biotin synthase in plants and seeds. The
CC   polypeptides may be used as a target to facilitate the design and/or
CC   identification of inhibitors of these enzymes that may be useful as
CC   herbicides. Furthermore, the polynucleotides may be used as probes for
CC   genetically and physically mapping the genes of which they are a part.
CC   The current sequence is that of the corn biotin synthase-related protein
CC   of the invention.
XX
SQ   Sequence 377 AA;

Alignment Scores:
Pred. No.:              1.52          Length:         377
Score:                  11.00         Matches:        11
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            2.1%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x ADE76386 (1-377)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 39
ADY09732
ID   ADY09732 standard; protein; 381 AA.
XX
AC   ADY09732;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polypeptide seqid 65547.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
```

PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 65547; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 381 AA;

Alignment Scores:

Pred. No.:	1.52	Length:	381
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADY09732 (1-381)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         23 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 33
```

RESULT 40
AAO30442
ID AAO30442 standard; protein; 631 AA.
XX
AC AAO30442;
XX
DT 22-SEP-2003 (first entry)
XX
DE Human MBM protein.
XX
KW Human; modifiers of branching morphogenesis; MBM; angiogenesis; cancer.
XX
OS Homo sapiens.
XX
PN WO2003045321-A2.
XX
PD 05-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US037777.
XX
PR 26-NOV-2001; 2001US-0333343P.
PR 26-NOV-2001; 2001US-0333371P.
PR 26-NOV-2001; 2001US-0333374P.
PR 26-NOV-2001; 2001US-0333379P.
PR 30-NOV-2001; 2001US-0334515P.
PR 30-NOV-2001; 2001US-0334537P.
PR 30-NOV-2001; 2001US-0334540P.
PR 30-NOV-2001; 2001US-0334542P.
PR 30-NOV-2001; 2001US-0334545P.
PR 05-DEC-2001; 2001US-0338276P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Karim FD, Keyes LN, Plowman GD;
XX
DR WPI; 2003-493375/46.
DR N-PSDB; AAL61274.
XX
PT Identifying modulators of branching morphogenesis (e.g. angiogenesis) for
PT use as therapeutic targets for cancers or for diagnosing cancers, by
PT detecting the presence of modifiers of branching morphogenesis (MBM)
PT proteins or genes.
XX
PS Example 2; Page 146-148; 169pp; English.
XX
CC The invention relates to a method of identifying a candidate branching
CC morphogenesis-modulating agent. The method involves detecting a test
CC agent-biased activity of the assay system which comprises modifiers of
CC branching morphogenesis (MBM) polypeptides or nucleic acids. The
CC difference between the test agent-biased activity and the reference
CC activity identifies the test agent as a candidate branching morphogenesis
CC modulating agent. The method is useful for identifying MBMs or modulators
CC of branching morphogenesis (specifically angiogenesis). The MBMs or
CC modulators are useful as therapeutic targets for disorders associated
CC with defective branching morphogenesis function or for diagnosing a
CC disease (particularly cancer). The present sequence is human MBM protein.
CC This sequence is used to illustrate the method of the invention
XX
SQ Sequence 631 AA;

Alignment Scores:

Pred. No.:	1.43	Length:	631
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x AAO30442 (1-631)

```
Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||
Db         92 SerSerGlyGlyGlySerSerGlySerSerSer 102
```

RESULT 41
ADJ70128
ID ADJ70128 standard; protein; 631 AA.
XX
AC ADJ70128;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID1934.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX

PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 1934; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 631 AA;

Alignment Scores:

Pred. No.:	1.43	Length:	631
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ADJ70128 (1-631)

```
Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||
Db         92 SerSerGlyGlyGlySerSerGlySerSerSer 102
```

RESULT 42
ADO41710
ID ADO41710 standard; protein; 651 AA.
XX
AC ADO41710;
XX
DT 15-JUL-2004 (first entry)
XX
DE Novel human polypeptide NOV28a.
XX
KW cardiant; antiarteriosclerotic; hypotensive; immunosuppressive;
KW dermatological; anorectic; cytostatic; antidiabetic; haemostatic;
KW anti-HIV; antiasthmatic; antibacterial; virucide; neuroprotective;
KW nootropic; antiparkinsonian; antilipemic; analgesic; antianginal;
KW gene therapy; vaccine; NOVX-associated disorder; cardiomyopathy;
KW atherosclerosis; hypertension; scleroderma; obesity; cancer; diabetes;
KW haemophilia; graft-versus-host disease; AIDS; asthma; Crohn's disease;
KW multiple sclerosis; infection; anorexia; cancer-associated cachexia;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW haematopoietic disorder; dyslipidemia; wasting disorder;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;
KW vaccine; human.
XX
OS Homo sapiens.
XX
PN US2004018555-A1.
XX
PD 29-JAN-2004.
XX
PF 03-JUN-2002; 2002US-00161493.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0297567P.
PR 12-JUN-2001; 2001US-0297573P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 15-JUN-2001; 2001US-0298556P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 12-SEP-2001; 2001US-0318771P.
PR 28-SEP-2001; 2001US-0325687P.
PR 24-OCT-2001; 2001US-0339266P.
PR 16-NOV-2001; 2001US-0337524P.
PR 14-DEC-2001; 2001US-0341143P.
PR 21-FEB-2002; 2002US-0358643P.
PR 21-FEB-2002; 2002US-0359151P.
PR 28-FEB-2002; 2002US-0361195P.
PR 05-MAR-2002; 2002US-0361964P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-APR-2002; 2002US-0371523P.
XX
PA (ANDE/) ANDERSON D W.
PA (ZERH/) ZERHUSEN B D.
PA (LILL/) LI L.
PA (ZHON/) ZHONG M.
PA (CASM/) CASMAN S J.
PA (GERL/) GERLACH V.
PA (SHIM/) SHIMKETS R A.
PA (GORM/) GORMAN L.
PA (PENA/) PENA C E A.
PA (KEKU/) KEKUDA R.
PA (PATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K A.
PA (LEIT/) LEITE M W.
PA (RAST/) RASTELLI L.
PA (MACD/) MACDOUGALL J R.
PA (TAUP/) TAUPIER R J.
PA (GUOX/) GUO X S.
PA (MILL/) MILLER C E.
PA (SHEN/) SHENOY S G.
PA (HJAL/) HJALT T.
PA (VOSS/) VOSS E Z.
PA (BOLD/) BOLDOG F L.
PA (MALY/) MALYANKAR U M.
PA (PADI/) PADIGARU M.
PA (JIWW/) JI W.
PA (SMIT/) SMITHSON G.
PA (EDIN/) EDINGER S R.
PA (MILL/) MILLET I.
PA (ELLE/) ELLERMAN K.
XX
PI Anderson DW, Zerhusen BD, Li L, Zhong M, Casman SJ, Gerlach V;

PI Shimkets RA, Gorman L, Pena CEA, Kekuda R, Patturajan M, Spytek KA;
PI Leite MW, Rastelli L, Macdougall JR, Taupier RJ, Guo XS, Miller CE;
PI Shenoy SG, Hjalt T, Voss EZ, Boldog FL, Malyankar UM, Padigaru M;
PI Ji W, Smithson G, Edinger SR, Millet I, Ellerman K;
XX
DR WPI; 2004-122030/12.
DR N-PSDB; ADO41709.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes,
PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
PT or pharmacogenomics.
XX
PS Claim 1; SEQ ID NO 76; 322pp; English.
XX
CC The invention describes a new polypeptide comprising: any of the 36 fully
CC defined amino acid sequences (e.g. 860, 919 or 681 amino acids) given in
CC the specification; a mature form of (a); a sequence that is at least 95%
CC identical to (a); or a sequence comprising one or more conservative
CC substitutions in (a). The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease, the
CC disease selected from a pathology associated with the polypeptide. These
CC may also be used in diagnosing, treating or preventing NOVX-associated
CC disorders such as cardiomyopathy, atherosclerosis, hypertension,
CC scleroderma, obesity, cancer, diabetes, haemophilia, graft-versus-host
CC disease, AIDS, asthma, Crohn's disease, multiple sclerosis, infections,
CC anorexia, cancer-associated cachexia, neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), haematopoietic disorders,
CC dyslipidemias and other wasting disorders associated with chronic
CC diseases. The nucleic acids are also used as hybridisation probes, in
CC chromosome mapping, tissue typing, preventive medicine, and
CC pharmacogenomics. The polypeptides are also useful as vaccines. This is
CC the amino acid sequence of a novel human polypeptide of the invention.
XX
SQ Sequence 651 AA;

Alignment Scores:

Pred. No.:	1.42	Length:	651
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADO41710 (1-651)

```
Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||||||
Db        170 SerSerGlyGlyGlySerSerGlySerSerSer 180
```

RESULT 43
ABJ19355
ID ABJ19355 standard; protein; 709 AA.
XX
AC ABJ19355;
XX
DT 28-MAR-2003 (first entry)
XX
DE NOVX related protein SEQ ID No 76.
XX
KW Antidiabetic; anorectic; virucide; antibacterial; fungicide; nootropic;
KW protozoacide; neuroprotective; antiparkinsonian; antilipaemic;
KW NOVX-associated disorder; metabolic disorder; diabetes; anorexia;
KW obesity; infectious disease; cancer-associated cachexia; immune disorder;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW haematopoietic disorder; cancer; dyslipidaemia; metabolic disturbance;
KW neurogenesis; cell differentiation; cell proliferation; haematopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; preventive medicine; pharmacogenomic; NOVX.
XX
OS Unidentified.
XX
PN WO200299062-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017559.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0297567P.
PR 12-JUN-2001; 2001US-0297573P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 15-JUN-2001; 2001US-0298556P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 12-SEP-2001; 2001US-0318771P.
PR 25-SEP-2001; 2001US-0324687P.
PR 24-OCT-2001; 2001US-0339266P.
PR 16-NOV-2001; 2001US-0337524P.
PR 14-DEC-2001; 2001US-0341143P.
PR 21-FEB-2002; 2002US-0358643P.
PR 21-FEB-2002; 2002US-0359151P.
PR 28-FEB-2002; 2002US-0361195P.
PR 05-MAR-2002; 2002US-0361964P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-APR-2002; 2002US-0371523P.
PR 03-JUN-2002; 2002US-00161493.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Zerhusen BD, Li L, Zhong M, Casman SJ, Gerlach VL;
PI Shimkets RA, Gorman L, Pena CEA, Kekuda R, Patturajan M, Spytek KA;
PI Leite MW, Rastelli L, Macdougall JR, Taupier RJ, Guo X, Miller CE;
PI Shenoy SG, Hjalt T, Voss EZ, Boldog FL, Malyankar UM, Padigaru M;
PI Ji W, Smithson G, Edinger SR, Millet I, Ellerman K;
XX
DR WPI; 2003-140607/13.
DR N-PSDB; ABT16054.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT obesity, cancer, Parkinson's disease, infections, immune disorders, or
PT various dyslipidemias.
XX
PS Example 28; Page 183; 461pp; English.
XX
CC The invention relates to an isolated polypeptide comprising any of the 36
CC 86-1370 residue amino acid sequences, given in the specification, a
CC mature form of them, or a sequence that is at least 95 % identical to, or
CC having one or more conservative amino acid substitutions in one of the 36
CC amino acid sequences. The polypeptides, nucleic acid molecules and
CC antibodies of the invention are useful in the manufacture of a medicament
CC for treating a syndrome associated with a human disease, preferably a
CC NOVX-associated disorder. The nucleic acid molecules, polypeptides and
CC antibodies are useful for treating, preventing or diagnosing diseases
CC such as metabolic disorders, diabetes, obesity, infectious diseases
CC (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer-
CC associated cachexia, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, immune disorders, haematopoietic disorders, cancer
CC and various dyslipidaemias, or metabolic disturbances associated with
CC obesity, metabolic X syndrome, and wasting disorders. The nucleic acids
CC and polypeptides may also be used as targets for the identification of
CC small molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVX substances for use in therapeutic or

```
CC   diagnostic methods. The nucleic acids are further used as hybridisation
CC   probes, in chromosome mapping, tissue typing, preventive medicine, and
CC   pharmacogenomics. This sequence represents a NOVX related protein of the
CC   invention
XX
SQ   Sequence 709 AA;

Alignment Scores:
Pred. No.:              1.4             Length:         709
Score:                  11.00           Matches:        11
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.1%            Indels:         0
DB:                     6               Gaps:           0

US-10-677-865A-2 (1-1590) x ABJ19355 (1-709)

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||||
Db        170 SerSerGlyGlyGlySerSerGlySerSerSer 180

RESULT 44
ADQ67680
ID   ADQ67680 standard; protein; 709 AA.
XX
AC   ADQ67680;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Novel human protein sequence #2346.
XX
KW   osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW   gene therapy; diagnostic marker; morbid state; osteoporosis;
KW   neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   EP1440981-A2.
XX
PD   28-JUL-2004.
XX
PF   21-JAN-2004; 2004EP-00001196.
XX
PR   21-JAN-2003; 2003JP-00102206.
PR   09-MAY-2003; 2003JP-00131392.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Nagai K,  Irie R;
XX
DR   WPI; 2004-535376/52.
DR   N-PSDB; ADQ67373.
XX
PT   Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT   Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS   Claim 1; SEQ ID NO 4841; 2449pp; English.
XX
CC   The invention relates to 2495 novel polynucleotides (I) and their encoded
CC   polypeptides, sequences hybridizing to these nucleotides, sequences
CC   encoding partial polypeptides and sequences having 70% or 90% identity to
CC   the nucleotide and protein sequences. The nucleotides and polypeptides
CC   are useful as diagnostic markers or therapeutic target for the diseases
CC   or morbid states. They are also useful for treating osteoporosis,
CC   neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC   dementia and various cancers. This sequence corresponds to a protein
CC   sequence of the invention.
XX
SQ   Sequence 709 AA;

Alignment Scores:
Pred. No.:              1.4             Length:         709
Score:                  11.00           Matches:        11
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.1%            Indels:         0
DB:                     8               Gaps:           0

US-10-677-865A-2 (1-1590) x ADQ67680 (1-709)

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||||
Db        170 SerSerGlyGlyGlySerSerGlySerSerSer 180

RESULT 45
ABB68240
ID   ABB68240 standard; protein; 828 AA.
XX
AC   ABB68240;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 31512.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL12343.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 31512; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 828 AA;

Alignment Scores:
Pred. No.:              1.38            Length:         828
Score:                  11.00           Matches:        11
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.1%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x ABB68240 (1-828)
```

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||||
Db        301 GlyGlySerSerGlySerSerSerGlySerSer 311
```

RESULT 46
ABG68619
ID ABG68619 standard; peptide; 10 AA.
XX
AC ABG68619;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human corneodesmosin peptide #1.
XX
KW Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS Homo sapiens.
XX
PN WO200244736-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-GB005303.
XX
PR 30-NOV-2000; 2000GB-00029225.
PR 07-DEC-2000; 2000GB-00029879.
XX
PA (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI Tazi-Ahnini R, Bavik C, Ward S, Duff G, Cork M;
XX
DR WPI; 2002-557554/59.
XX
PT Diagnosing a disease or susceptibility to a disease associated with
PT abnormal cell-cell adhesion between epithelial cells, by detecting
PT mutation in nucleic acid encoding adhesion protein, protease or protease
PT inhibitor.
XX
PS Example A5; Page 101; 257pp; English.
XX
CC The invention relates to a method for diagnosis of a disease or
CC susceptibility to a disease associated with abnormal cell-cell adhesion
CC between epithelial cells, comprising detecting a mutation in a nucleic
CC acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC modulated level of adhesion protein, protease or protease inhibitor, or
CC its fragment polypeptide. The method is useful for diagnosing a disease
CC such as eczema, contact dermatitis, lung atopic asthma, post viral
CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC malignancy of the gastrointestinal tract, malignancy of the lung,
CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC a peptide used in the scope of the invention
XX
SQ Sequence 10 AA;

Alignment Scores:

Pred. No.:	19	Length:	10
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x ABG68619 (1-10)

```
Qy        142 ATCACCTCCCCTAACGACCCCTGCCTCACT 171
              ||||||||||||||||||||||||||||||
Db          1 IleThrSerProAsnAspProCysLeuThr 10
```

RESULT 47
ADN19766
ID ADN19766 standard; protein; 309 AA.
XX
AC ADN19766;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #2419.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 2419; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 309 AA;

Alignment Scores:

Pred. No.:	12.4	Length:	309
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADN19766 (1-309)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        150 SerSerSerHisSerSerSerSerSerSer 159
```

RESULT 48
ADN19767
ID ADN19767 standard; protein; 344 AA.
XX
AC ADN19767;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #2420.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 2420; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 344 AA;

Alignment Scores:

Pred. No.:	12.2	Length:	344
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	8	Gaps:	0

US-10-677-865A-2 (1-1590) x ADN19767 (1-344)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        185 SerSerSerHisSerSerSerSerSerSer 194
```

RESULT 49
AAR54665
ID AAR54665 standard; protein; 348 AA.
XX
AC AAR54665;
XX
DT 25-MAR-2003 (revised)
DT 06-NOV-1994 (first entry)
XX
DE Outer membrane lipoprotein A.
XX
KW OMP; Actinobacillus pleuropneumoniae; AP; vaccine; pigs; probe.
XX
OS Actinobacillus pleuropneumoniae.
XX
PN WO9410316-A1.
XX
PD 11-MAY-1994.
XX
PF 03-NOV-1993; 93WO-CA000448.
XX
PR 05-NOV-1992; 92US-00971558.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Gerlach GF, Willson PJ, Rossi-Campos A, Potter AA;
XX
DR WPI; 1994-167471/20.
DR N-PSDB; AAQ65353.
XX
PT Pure actinobacillus pleuropneumoniae outer membrane lipoprotein - used in
PT vaccines for treating or preventing infection, partic. in pigs and for
PT diagnosis.
XX
PS Claim 1; Fig 1; 59pp; English.
XX
CC Actinobacillus pleuropneumoniae was cultured and outer membrane protein
CC (OMPs) were isolated and analysed. AP culture supernatant was used to
CC raise rabbit antiserum which was used to screen an AP strain AP 37
CC genomic DNA library in E. coli to obtain DNA encoding AP serotype 1 OmlA.
CC The protein prod. of the gene can be used in avccines for treating AP in
CC vertebrates, esp. pigs. The gene may be used as a probe for detecting AP

```
CC   and for detecting homologous genes in other bacterial strains. See also
CC   AAR54666. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 348 AA;

Alignment Scores:
Pred. No.:              12.2          Length:         348
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x AAR54665 (1-348)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30

RESULT 50
ADB65137
ID   ADB65137 standard; protein; 406 AA.
XX
AC   ADB65137;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human protein encoded by clone SPLEN20063250.
XX
KW   Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW   cell regeneration; membrane protein; signal transduction-related protein;
KW   transcription-related protein; osteoporosis; neurological disease;
KW   cancer; tumour.
XX
OS   Homo sapiens.
XX
PN   EP1308459-A2.
XX
PD   07-MAY-2003.
XX
PF   28-MAR-2002; 2002EP-00007401.
XX
PR   05-NOV-2001; 2001JP-00379298.
PR   25-JAN-2002; 2002US-00350978.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-450961/43.
DR   N-PSDB; ADB63167.
XX
PT   New polynucleotides and polypeptides, useful for developing a diagnostic
PT   marker or medicines for regulation of their expression and activity, or
PT   as targets of gene therapy.
XX
PS   Claim 1; Page; 222pp; English.
XX
CC   The invention discloses a polynucleotide comprising a sequence selected
CC   from 1970 fully defined nucleotide sequences which encode novel
CC   polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC   or its partial peptide, an antibody binding to the polypeptide or peptide
CC   of the polynucleotide, immunologically assaying the polypeptide or
CC   peptide of the polynucleotide by contacting the polypeptide or peptide
CC   with the antibody of the encoded protein, and observing the binding
CC   between the two, a transformant carrying the polynucleotide in an
CC   expressible manner and an antisense polynucleotide. The oligonucleotide
CC   is useful as a primer for synthesising the polynucleotide, or as a probe
CC   for detecting the polynucleotide. The polynucleotides and encoded
CC   proteins are useful as pharmaceutical agents and many disease-related
CC   genes may be included in them, for developing a diagnostic marker or
CC   medicines for regulation of their expression and activity, or as targets
CC   of gene therapy. The genes are involved in tissue and/or cell
CC   regeneration. Membrane proteins, signal transduction-related proteins,
CC   transcription-related proteins, disease-related proteins and genes
CC   encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC   neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC   the activity or expression of the encoded protein to treat diseases. The
CC   sequence presented is a protein of the invention. Note: Some of the
CC   sequence data for this patent is not represented in the printed
CC   specification, but is based on sequence information supplied by the
CC   European Patent Office.
XX
SQ   Sequence 406 AA;

Alignment Scores:
Pred. No.:              12            Length:         406
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x ADB65137 (1-406)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81

RESULT 51
ADR09419
ID   ADR09419 standard; protein; 406 AA.
XX
AC   ADR09419;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Human protein useful for treating neurological disease Seq 2925.
XX
KW   human; oligo-capping method; diagnostic marker; gene therapy;
KW   osteoporosis; neurological disease; Alzheimer's disease;
KW   Parkinson's disease; dementia; short memory; cancer;
KW   sense or motor function; emotional reaction; fear response; panic;
KW   osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW   tranquiliser.
XX
OS   Homo sapiens.
XX
PN   EP1447413-A2.
XX
PD   18-AUG-2004.
XX
PF   12-FEB-2004; 2004EP-00003145.
XX
PR   14-FEB-2003; 2003JP-00102207.
PR   09-MAY-2003; 2003JP-00131452.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Yamamoto J,  Nishikawa T,  Isono Y,  Sugiyama T,  Otsuki T;
PI   Wakamatsu A,  Ishii S,  Nagai K,  Irie R;
XX
DR   WPI; 2004-583265/57.
DR   N-PSDB; ADR07463.
XX
PT   New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT   Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS   Claim 1; SEQ ID NO 2925; 2686pp; English.
XX
CC   This invention relates to novel, isolated full length human cDNA
CC   molecules and the encoded proteins thereof. Specifically, it refers to
CC   cDNA clones obtained by an oligo-capping method, where none of these
```

```
CC   clones are identical to any known human mRNAs. The present invention
CC   describes an immunoassay to identify agonists and antagonists, as well as
CC   antibodies, antisense molecules and siRNAs that can all be used to bind
CC   to and modulate expression of the cDNA molecules. As such, these
CC   molecules are useful for diagnostic markers or therapeutic targets for
CC   the various diseases or morbid states. In particular, they are useful in
CC   gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC   disease, Parkinson's disease, dementia, short memory and various cancers,
CC   as well as for maintaining equilibrium of sense or motor function, and
CC   for treating emotional reaction, fear response and panic. Accordingly,
CC   they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC   cytostatic and tranquiliser activities. This polypeptide is a protein
CC   encoded by a full length human cDNA sequence of the invention. NOTE: This
CC   sequence is not given in the sequence listing of the specification but
CC   can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC   office.
XX
SQ   Sequence 406 AA;

Alignment Scores:
Pred. No.:              12              Length:         406
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     8               Gaps:           0

US-10-677-865A-2 (1-1590) x ADR09419 (1-406)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81

RESULT 52
AAW37241
ID   AAW37241 standard; protein; 476 AA.
XX
AC   AAW37241;
XX
DT   20-JUL-1998  (first entry)
XX
DE   Heterodera glycines cellulase HG-eng1.
XX
KW   Cellulase; endoglucanase-1; Hg-eng1; nematode; crop protection;
KW   disease resistance; pest resistance; cellulose; feed additive; food;
KW   waste water.
XX
OS   Heterodera glycines.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /label= Sig_peptide
FT   Protein         29. .476
FT                   /label= Mat_protein
FT   Domain          29. .330
FT                   /note= "catalytic domain"
FT   Peptide         331. .380
FT                   /note= "linker sequence"
FT   Domain          381. .476
FT                   /note= "cellulose binding domain"
XX
PN   WO9801569-A1.
XX
PD   15-JAN-1998.
XX
PF   08-JUL-1997;   97WO-NL000400.
XX
PR   08-JUL-1996;   96EP-00201890.
PR   17-JAN-1997;   97EP-00200136.
XX
PA   (RIJK-) RIJKSLANDBOUWHOGESCHOOL.
XX
PI   Schots A,  Bakker J,  Helder J,  Gommers F,  Stiekema WJ,  Roosien J;
PI   Goverse A,  Schouten A,  Smant G,  De Boer J,  Stokkermans JPW,  Abad P;
PI   Rosso MNF;
XX
DR   WPI; 1998-101064/09.
DR   N-PSDB; AAV03026.
XX
PT   Cellulases from nematodes and related nucleic acid, vectors, transformed
PT   cells and antibodies - useful for protecting plants against parasitic
PT   nematodes, for treating cellulosic fibres, as food or feed additive, in
PT   waste water treatment etc.
XX
PS   Claim 1; Page 17-19; 71pp; English.
XX
CC   This polypeptide comprises a cellulase, designated HG-eng1, from the
CC   salivary glands of the nematode Heterodera glycines. Its amino acid
CC   sequence was deduced from an isolated cDNA clone (see AAV03026). The
CC   invention relates to novel nematode peptides having cellulase activity
CC   (see AAW37240-46) and nucleic acids encoding them (see AAV03025-31 and
CC   AAT99672-74). It also provides expression vectors containing such nucleic
CC   acids, recombinant host cells, and antibodies against the cellulases. The
CC   enzymic peptides are useful in the protection of plants against parasitic
CC   nematodes, and can also be used in the feed, animal feed and bevarage
CC   industries, in the paper and clothing industries, and in waste-water
CC   treatment. The antibodies can also be used to protect plants, by
CC   inhibiting parasitic cellulases
XX
SQ   Sequence 476 AA;

Alignment Scores:
Pred. No.:              11.7            Length:         476
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     2               Gaps:           0

US-10-677-865A-2 (1-1590) x AAW37241 (1-476)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        344 GlySerSerGlySerSerSerGlySerSer 353

RESULT 53
AAW43910
ID   AAW43910 standard; protein; 484 AA.
XX
AC   AAW43910;
XX
DT   25-MAR-2003  (revised)
DT   17-JUN-1998  (first entry)
XX
DE   Heterodera glycines svp encoded cellulase.
XX
KW   Cellulase activity; nematode; crop protection; parasitic;
KW   subventral oesophageal protein.
XX
OS   Heterodera glycines.
XX
PN   EP818538-A1.
XX
PD   14-JAN-1998.
XX
PF   08-JUL-1996;   96EP-00201890.
XX
PR   08-JUL-1996;   96EP-00201890.
XX
PA   (RIJK-) RIJKSLANDBOUWHOGESCHOOL.
XX
PI   Schots A,  Bakker J,  Helder J,  Gommers FJ,  Stiekema WJ,  Roosien J;
PI   Goverse A,  Schouten A,  Smant G,  De Boer JM,  Stokkermans JPW;
XX
DR   WPI; 1998-065307/07.
DR   N-PSDB; AAT86645.
```

XX
PT Nematodal cellulase peptide(s) - useful as industrial cellulolytic
PT enzymes or for plant protection.
XX
PS Claim 1; Page 18-20; 28pp; English.
XX
CC The sequence is that of a peptide having cellulase activity. The peptide
CC comprises at least a fragment of nematode subventral oesophageal proteins
CC (svp's), can be used as a cellulolytic enzyme in the food and beverage
CC industry, in the paper or clothing industry, or in waste treatment. The
CC peptide or antibodies directed against it can be used to protect plants
CC against parasitic nematodes. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
SQ Sequence 484 AA;

Alignment Scores:

Pred. No.:	11.7	Length:	484
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x AAW43910 (1-484)

```
Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        315 GlySerSerGlySerSerSerGlySerSer 324
```

RESULT 54
AAO20499
ID AAO20499 standard; protein; 593 AA.
XX
AC AAO20499;
XX
DT 27-JUN-2002 (first entry)
XX
DE Protein of APP related human homologue hCP41313 #1.
XX
KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;
KW APP pathway modulator; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200226820-A2.
XX
PD 04-APR-2002.
XX
PF 01-OCT-2001; 2001WO-EP011345.
XX
PR 29-SEP-2000; 2000US-0236893P.
PR 14-JUN-2001; 2001US-0298309P.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
PI Reinhardt MWHM, Zusman S;
XX
DR WPI; 2002-315796/35.
DR N-PSDB; AAK99393.
XX
PT New transgenic fly, containing DNA encoding an Abeta portion of human
PT APP, useful for identifying agents which modulate the APP pathway and
PT which can be used to treat Alzheimer's disease.
XX
PS Claim 25; Page 90-91; 129pp; English.
XX
CC The invention relates to a transgenic fly whose genome comprises DNA
CC encoding a polypeptide having the Abeta portion of human amyloid
CC precursor protein (APP), fused to a signal sequence. The DNA sequence
CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC the specification. The DNA sequence is operably linked to a tissue-
CC specific expression control sequence. Expression of the sequence gives
CC the fly an altered phenotype. The purpose of the invention is for
CC identifying agents that inhibit or promote the expression and/or function
CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This sequence represents the
CC protein of the APP related human homologue hCP41313 #1
XX
SQ Sequence 593 AA;

Alignment Scores:

Pred. No.:	11.4	Length:	593
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAO20499 (1-593)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        261 GlyGlySerSerGlySerSerSerGlySer 270
```

RESULT 55
ADJ70515
ID ADJ70515 standard; protein; 609 AA.
XX
AC ADJ70515;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID2321.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 2321; 180pp; English.
XX

```
CC   This invention relates to novel mitochondrial targets that can be used
CC   for therapeutic intervention in treating a disease associated with
CC   altered mitochondrial function. Specifically, it refers to a method for
CC   identifying proteins of the human heart mitochondrial proteome that are
CC   useful for drug screening assays, as well as therapeutic targets. The
CC   present invention describes a method for identifying such proteins that
CC   can be used in the treatment of various diseases associated with altered
CC   mitochondrial function including diabetes mellitus, Huntington's disease,
CC   osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC   encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC   ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC   compositions have neuroprotective, nootropic, antidiabetic,
CC   anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC   cytostatic activities. This polypeptide sequence is a human heart
CC   mitochondrial protein of the invention.
XX
SQ   Sequence 609 AA;

Alignment Scores:
Pred. No.:              11.4          Length:         609
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x ADJ70515 (1-609)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        275 GlyGlySerSerGlySerSerSerGlySer 284

RESULT 56
ADZ80776
ID   ADZ80776 standard; protein; 619 AA.
XX
AC   ADZ80776;
XX
DT   14-JUL-2005  (first entry)
XX
DE   Amino acid sequence of human homologue of Drosophila gene CG4220.
XX
KW   neurodegenerative disease; nootropic; neuroprotective; gene therapy;
KW   Alzheimer's Disease; degeneration; neurological disease; CG4220.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "encoded by TA"
XX
PN   WO2005039635-A2.
XX
PD   06-MAY-2005.
XX
PF   21-OCT-2004; 2004WO-EP011926.
XX
PR   22-OCT-2003; 2003US-0513241P.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS PHARMA GMBH.
XX
PI   Cohen D,  Dengler UJ,  Garza D,  Kelkar AN,  Konsolaki M,  Song H;
XX
DR   WPI; 2005-322916/33.
DR   N-PSDB; ADZ80775.
XX
PT   Treating, preventing or ameliorating neurodegenerative conditions, e.g.
PT   Alzheimer's disease, comprises administering to the subject a modulator
PT   of a protein, e.g. proteins involved in neurodegenerative conditions.
XX
PS   Claim 1; SEQ ID NO 4; 102pp; English.
XX
CC   The specification describes a method for treating, preventing or
CC   ameliorating neurodegenerative conditions. The method comprises
CC   administering to the subject an effective amount of a modulator of a
CC   protein, which is selected from the human homologs of several Drosophila
CC   genes. The method and modulators are useful for treating, preventing or
CC   ameliorating neurodegenerative conditions, e.g. Alzheimer's Disease. The
CC   present sequence is encoded by the human homologue of Drosophila gene
CC   CG4220, whose modulators can be used in the method of the invention.
XX
SQ   Sequence 619 AA;

Alignment Scores:
Pred. No.:              11.4          Length:         619
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     9             Gaps:           0

US-10-677-865A-2 (1-1590) x ADZ80776 (1-619)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        285 GlyGlySerSerGlySerSerSerGlySer 294

RESULT 57
AAO20500
ID   AAO20500 standard; protein; 645 AA.
XX
AC   AAO20500;
XX
DT   27-JUN-2002  (first entry)
XX
DE   Protein of APP related human homologue hCP41313 #2.
XX
KW   Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW   amyloid precursor protein; tissue-specific expression control; human APP;
KW   APP pathway modulator; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200226820-A2.
XX
PD   04-APR-2002.
XX
PF   01-OCT-2001; 2001WO-EP011345.
XX
PR   29-SEP-2000; 2000US-0236893P.
PR   14-JUN-2001; 2001US-0298309P.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   Cohen D,  Dengler UJ,  Finelli AL,  Freuler F,  Konsolaki M;
PI   Reinhardt MWHM,  Zusman S;
XX
DR   WPI; 2002-315796/35.
DR   N-PSDB; AAK99394.
XX
PT   New transgenic fly, containing DNA encoding an Abeta portion of human
PT   APP, useful for identifying agents which modulate the APP pathway and
PT   which can be used to treat Alzheimer's disease.
XX
PS   Claim 25; Page 92-93; 129pp; English.
XX
CC   The invention relates to a transgenic fly whose genome comprises DNA
CC   encoding a polypeptide having the Abeta portion of human amyloid
CC   precursor protein (APP), fused to a signal sequence. The DNA sequence
CC   encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC   the specification. The DNA sequence is operably linked to a tissue-
CC   specific expression control sequence. Expression of the sequence gives
CC   the fly an altered phenotype. The purpose of the invention is for
CC   identifying agents that inhibit or promote the expression and/or function
```

CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This sequence represents the
CC protein of the APP related human homologue hCP41313 #2
XX
SQ Sequence 645 AA;

Alignment Scores:

Pred. No.:	11.3	Length:	645
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAO20500 (1-645)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        311 GlyGlySerSerGlySerSerSerGlySer 320
```

RESULT 58
AAO20518
ID AAO20518 standard; protein; 646 AA.
XX
AC AAO20518;
XX
DT 27-JUN-2002 (first entry)
XX
DE Protein of APP related human homologue hCP41313 #3.
XX
KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;
KW APP pathway modulator; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200226820-A2.
XX
PD 04-APR-2002.
XX
PF 01-OCT-2001; 2001WO-EP011345.
XX
PR 29-SEP-2000; 2000US-0236893P.
PR 14-JUN-2001; 2001US-0298309P.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
PI Reinhardt MWHM, Zusman S;
XX
DR WPI; 2002-315796/35.
DR N-PSDB; AAK99412.
XX
PT New transgenic fly, containing DNA encoding an Abeta portion of human
PT APP, useful for identifying agents which modulate the APP pathway and
PT which can be used to treat Alzheimer's disease.
XX
PS Claim 25; Page 128-129; 129pp; English.
XX
CC The invention relates to a transgenic fly whose genome comprises DNA
CC encoding a polypeptide having the Abeta portion of human amyloid
CC precursor protein (APP), fused to a signal sequence. The DNA sequence
CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC the specification. The DNA sequence is operably linked to a tissue-
CC specific expression control sequence. Expression of the sequence gives
CC the fly an altered phenotype. The purpose of the invention is for
CC identifying agents that inhibit or promote the expression and/or function
CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This sequence represents the
CC protein of the APP related human homologue hCP41313 #3
XX
SQ Sequence 646 AA;

Alignment Scores:

Pred. No.:	11.3	Length:	646
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x AAO20518 (1-646)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        312 GlyGlySerSerGlySerSerSerGlySer 321
```

RESULT 59
ABG74687
ID ABG74687 standard; protein; 651 AA.
XX
AC ABG74687;
XX
DT 10-MAY-2003 (first entry)
XX
DE Human CGDD protein 914113CD1 SEQ ID 13.
XX
KW Human; cell growth, differentiation and death; CGDD; cardiant; cancer;
KW cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
KW neurological disorder.
XX
OS Homo sapiens.
XX
PN WO2003014322-A2.
XX
PD 20-FEB-2003.
XX
PF 08-AUG-2002; 2002WO-US025465.
XX
PR 08-AUG-2001; 2001US-0311017P.
PR 17-AUG-2001; 2001US-0313070P.
PR 17-AUG-2001; 2001US-0313071P.
PR 24-AUG-2001; 2001US-0314678P.
PR 31-AUG-2001; 2001US-0316692P.
PR 07-SEP-2001; 2001US-0317913P.
PR 14-SEP-2001; 2001US-0322182P.
PR 07-DEC-2001; 2001US-0340747P.
PR 20-DEC-2001; 2001US-0342761P.
PR 29-MAR-2002; 2002US-0369129P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;
PI Duggan BM, Elliott VS, Emerling BM, Forsythe IJ, Gietzen KJ;
PI Gorvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
PI Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
PI Yue H;
XX
DR WPI; 2003-268197/26.
DR N-PSDB; ABQ77426.
XX
PT New polypeptide associated with cell growth, differentiation and death,
PT useful for preparing a composition for diagnosing or treating a disease
PT e.g., cardiovascular or neurological disorders or cancer.
XX
PS Claim 68; Page 216-218; 267pp; English.
XX
CC This invention describes a novel polypeptide associated with cell growth,

CC differentiation and death (CGDD) which has cardiant, cytostatic and
CC neuroprotective activity. The polypeptides of the invention are useful
CC for preparing a composition for diagnosing, or treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional CGDD e.g., cardiovascular or neurological disorders or cancer.
CC The polynucleotides of the invention can be used for gene therapy.
CC ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
CC ABQ77414-ABQ77440
XX
SQ Sequence 651 AA;

Alignment Scores:

Pred. No.:	11.3	Length:	651
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x ABG74687 (1-651)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        317 GlyGlySerSerGlySerSerSerGlySer 326
```

RESULT 60
ABU03140
ID ABU03140 standard; protein; 866 AA.
XX
AC ABU03140;
XX
DT 21-JAN-2003 (first entry)
XX
DE Alpha amylase polypeptide #101.
XX
KW Alpha amylase; enzyme; starch linkage hydrolysis; starch liquefaction;
KW starch breakdown catalysis; textile desizing; lignocellulosic fibre;
KW enzymatic de-inking; recycled paper; high-maltose syrup; dough;
KW high glucose syrup; corn-wet milling process; detergent; baking process;
KW beverage; oil field; fuel ethanol; brewing process; staling;
KW starch modification.
XX
OS Unidentified.
XX
PN WO200268589-A2.
XX
PD 06-SEP-2002.
XX
PF 21-FEB-2002; 2002WO-US005068.
XX
PR 21-FEB-2001; 2001US-0270495P.
PR 21-FEB-2001; 2001US-0270496P.
PR 14-MAY-2001; 2001US-0291122P.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Callen W, Richardson T, Frey G;
XX
DR WPI; 2003-018656/01.
DR N-PSDB; ABX08511.
XX
PT Novel purified polypeptide with alpha-amylase activity, useful e.g. for
PT liquefying starch, for textile desizing, for treating lignocellulosic
PT fibers, and for producing high-maltose or high-glucose syrup.
XX
PS Claim 30; Fig 16; 301pp; English.
XX
CC The invention relates to a purified polypeptide with alpha-amylase
CC activity and the polynucleotide encoding it. The polypeptide is useful
CC for hydrolysing starch linkages, for catalysing the breakdown of a
CC starch, for modifying small molecules, for liquefying starch, for washing
CC an object, for textile desizing, for treating lignocellulosic fibers, for
CC improving fibre properties, for enzymatic de-inking of recycled paper
CC pulp, for producing a high-maltose or high-glucose syrup or a mixed
CC syrup, and for increasing the flow of production fluids from a
CC subterranean formation by removing a viscous, starch-containing, damaging
CC fluid formed during production operations and found within the
CC subterranean formation which surrounds a completed well bore. The
CC polypeptide is also useful for preparing a dough or a baked product
CC prepared from the dough and in corn-wet milling processes, detergents,
CC baking processes, beverages, oil fields (fuel ethanol), brewing processes
CC and starch modification in the paper and pulp industry, for removing
CC starch containing stains from a material and for reducing staling of
CC bakery products. Sequences ABU03040-ABU03144 represent alpha amylase
CC polypeptides of the invention
XX
SQ Sequence 866 AA;

Alignment Scores:

Pred. No.:	10.9	Length:	866
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x ABU03140 (1-866)

```
Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        462 GlySerSerGlySerSerSerGlySerSer 471
```

RESULT 61
ADU07494
ID ADU07494 standard; protein; 866 AA.
XX
AC ADU07494;
XX
DT 13-JAN-2005 (first entry)
XX
DE Amino acid sequence for amylase #101.
XX
KW Amylase; glucoamylase; 1,4-alpha-D-glucan glycohydrolase; alpha-amylase;
KW exoamylase; beta-amylase; glucosidic bond hydrolysis; starch; sugar;
KW glucose; maltodextrin; thermostable; enzyme.
XX
OS Unidentified.
XX
PN WO2004091544-A2.
XX
PD 28-OCT-2004.
XX
PF 08-MAR-2004; 2004WO-US007096.
XX
PR 06-MAR-2003; 2003US-00385305.
PR 28-MAR-2003; 2003US-0459014P.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Callen W, Richardson T, Frey G, Gray K, Kerovuo JS, Slupska M;
PI Barton N, O'donoghue E, Miller C;
XX
DR WPI; 2004-775532/76.
DR N-PSDB; ADU07493.
XX
PT Novel recombinant amylases, useful for hydrolysis of starch into sugars,
PT in detergent composition, for producing high-maltose or high-glucose
PT syrup, in brewing or alcohol production, bio-bleaching composition of
PT paper or pulp product.
XX
PS Claim 55; SEQ ID NO 204; 436pp; English.
XX
CC The invention relates to the isolation of polypeptides with amylase
CC activity, and polynucleotide sequences encoding the polypeptides. The
CC amylase activity comprises a glucoamylase activity, 1,4-alpha-D-glucan
CC glycohydrolase activity, alpha-amylase activity, exoamylase activity or

```
CC   beta-amylase activity. The amylase activity comprises hydrolyzing
CC   glucosidic bonds in a starch to produce sugars such as glucose and
CC   maltodextrins. The amylase activity comprises cleaving a maltose or a D-
CC   glucose unit from the non-reducing end of the starch. The amylase
CC   activity is thermostable and the polypeptides of the invention are active
CC   over a wide range of temperatures. The polypeptide sequences are useful
CC   for making anti-amylase antibodies, for identifying a modulator of an
CC   amylase activity, or identifying an amylase substrate. The polynucleotide
CC   sequences are useful for producing recombinant polypeptides, and for
CC   modifying codons in a nucleic acid encoding a polypeptide with an amylase
CC   activity to increase or decrease its expression in a host cell. The
CC   polypeptide and polynucleotide sequences are useful for making or
CC   modifying a small molecule, and for determining a functional fragment of
CC   an amylase enzyme. Primers to the polynucleotides are useful for
CC   amplifying them. This sequence represents a polypeptide sequence of the
CC   invention.
XX
SQ   Sequence 866 AA;

Alignment Scores:
Pred. No.:              10.9          Length:         866
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     8             Gaps:           0

US-10-677-865A-2 (1-1590) x ADU07494 (1-866)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        462 GlySerSerGlySerSerSerGlySerSer 471

RESULT 62
ABB65772
ID   ABB65772 standard; protein; 5533 AA.
XX
AC   ABB65772;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 24108.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL09875.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 24108; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5533 AA;

Alignment Scores:
Pred. No.:              8.63          Length:         5533
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x ABB65772 (1-5533)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 63
ABB71160
ID   ABB71160 standard; protein; 5560 AA.
XX
AC   ABB71160;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 40272.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL15263.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 40272; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5560 AA;
```

```
Alignment Scores:
Pred. No.:              8.62          Length:          5560
Score:                  10.00         Matches:         10
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.9%          Indels:          0
DB:                     4             Gaps:            0

US-10-677-865A-2 (1-1590) x ABB71160 (1-5560)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 64
ADO01004
ID   ADO01004 standard; protein; 5560 AA.
XX
AC   ADO01004;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Fruit fly AD-related protein CG18497 (Split ends).
XX
KW   Fruit fly; Alzheimer's disease; Gamma secretase; Psn gene; P-element; EP;
KW   APPL-SV; Amyloid precursor-like protein; APP;
KW   suppressor of hairless transcription factor; Su(H);
KW   VP16 activation domain; dementia; memory loss; language deterioration;
KW   impaired visuospatial skill.
XX
OS   Drosophila melanogaster.
XX
PN   US2004067535-A1.
XX
PD   08-APR-2004.
XX
PF   03-OCT-2002; 2002US-00263929.
XX
PR   03-OCT-2002; 2002US-00263929.
XX
PA   (LIFE-) LIFE SCI DEV CORP.
XX
PI   Kim J,  Galant R;
XX
DR   WPI; 2004-355296/33.
DR   N-PSDB; ADO00902.
XX
PT   Identifying compound by exposing cell that expresses gene having
PT   enhancing or suppression effect on APPL-SV phenotype to agent,
PT   identifying modulation of Alzheimer's disease (AD), regulation of gene or
PT   protein expression with AD.
XX
PS   Claim 18; SEQ ID NO 142; 185pp; English.
XX
CC   The invention relates to identifying a compound comprising exposing cell
CC   expressing gene 1 having enhancing or suppression effect on an APPL-SV
CC   phenotype (a transgenic fruit fly expressing the Amyloid precursor-like
CC   protein, APP, as a fusion protein with the suppressor of hairless
CC   transcription factor , Su(H) and VP16 activation domain. The fusion
CC   protein is cleaved by gamma secretase (encoded by the Psn gene) to
CC   release the Su(H-VP16, which affects wing vein development. Genes
CC   affecting Psn expression/activity were screened by crossing the APP-SV
CC   line with an EP P-element insertion library, and the DNA recovered from
CC   the appropriate EP strain and sequenced) chosen from ADO00863-ADO00964,
CC   being the identified fruit fly genes affecting APP processing and their
CC   mammalian homologues, identifying modulation of Alzheimer's disease (AD)
CC   symptom, regulation of biological pathway, gene expression or protein
CC   function associated with AD relative to cell in absence of agent. Also
CC   included are regulating AD (involves providing a subject with AD or
CC   symptoms of AD and an agent that changes the expression of a gene
CC   detailed above or changes the activity of a polypeptide having a sequence
CC   chosen from ADO00965-ADO01066, and treating the subject with the agent)
CC   and a composition (comprising a nucleic acid encoding a polypeptide
CC   detailed above or an expression vector comprising the nucleic acid or a
CC   host cell comprising the expression vector or an antisense
CC   oligonucleotide that hybridises under stringent conditions to the nucleic
CC   acid or polypeptide or an antibody that specifically binds to the
CC   polypeptide). The method is useful for identifying compounds modulating
CC   symptom of Alzheimer's disease (AD), regulation of biological pathway
CC   associated with AD, or regulation of gene expression or protein function
CC   of gene or protein associated with AD. The nucleic acids and proteins are
CC   useful in drug screening and useful in screening and treating the subject
CC   having increased susceptibility to AD or symptoms of AD such as dementia,
CC   memory loss, language deterioration and impaired visuospatial skills. The
CC   present sequence is a fruit fly protein from a gene identified as having
CC   an effect on the APP-SV phenotype.
XX
SQ   Sequence 5560 AA;

Alignment Scores:
Pred. No.:              8.62          Length:          5560
Score:                  10.00         Matches:         10
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.9%          Indels:          0
DB:                     8             Gaps:            0

US-10-677-865A-2 (1-1590) x ADO01004 (1-5560)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 65
ABG68625
ID   ABG68625 standard; peptide; 9 AA.
XX
AC   ABG68625;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Corneodesmosin peptide #6.
XX
KW   Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW   stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW   desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW   contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW   bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW   Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW   impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW   basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW   gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS   Unidentified.
XX
PN   WO200244736-A2.
XX
PD   06-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-GB005303.
XX
PR   30-NOV-2000; 2000GB-00029225.
PR   07-DEC-2000; 2000GB-00029879.
XX
PA   (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI   Tazi-Ahnini R,  Bavik C,  Ward S,  Duff G,  Cork M;
XX
DR   WPI; 2002-557554/59.
XX
PT   Diagnosing a disease or susceptibility to a disease associated with
PT   abnormal cell-cell adhesion between epithelial cells, by detecting
PT   mutation in nucleic acid encoding adhesion protein, protease or protease
PT   inhibitor.
```

XX
PS Example A5; Page 102; 257pp; English.
XX
CC The invention relates to a method for diagnosis of a disease or
CC susceptibility to a disease associated with abnormal cell-cell adhesion
CC between epithelial cells, comprising detecting a mutation in a nucleic
CC acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC modulated level of adhesion protein, protease or protease inhibitor, or
CC its fragment polypeptide. The method is useful for diagnosing a disease
CC such as eczema, contact dermatitis, lung atopic asthma, post viral
CC asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC malignancy of the gastrointestinal tract, malignancy of the lung,
CC psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC a peptide used in the scope of the invention
XX
SQ Sequence 9 AA;

Alignment Scores:

Pred. No.:	9.69e+07	Length:	9
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x ABG68625 (1-9)

```
Qy        169 ACTGGGAAGGGTGACTCCAGCGGCTTC 195
              |||||||||||||||||||||||||||
Db          1 ThrGlyLysGlyAspSerSerGlyPhe 9
```

RESULT 66
ABJ39151
ID ABJ39151 standard; peptide; 9 AA.
XX
AC ABJ39151;
XX
DT 06-NOV-2003 (first entry)
XX
DE Serine rich domain 2 of the human S protein (skin corneodesmosin).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC the dermatological activity of CDSN provides a good diagnostic and
CC therapeutic target for alopecia, and through gene therapy CDSN can be
CC used to modulate hair growth and treat male pattern baldness. This
CC peptide sequence is a functional region of the human S protein, the
CC cornesodesmosin protein solely expressed in the skin that is a target of
CC the invention
XX
SQ Sequence 9 AA;

Alignment Scores:

Pred. No.:	9.69e+07	Length:	9
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x ABJ39151 (1-9)

```
Qy        406 AGCAGCAGCTCTCACTCGGGAAGCAGC 432
              |||||||||||||||||||||||||||
Db          1 SerSerSerSerHisSerGlySerSer 9
```

RESULT 67
AAG94132
ID AAG94132 standard; peptide; 10 AA.
XX
AC AAG94132;
XX
DT 18-SEP-2001 (first entry)
XX
DE Human complementary peptide, SEQ ID NO: 326.
XX
KW Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS Homo sapiens.
XX
PN WO200142277-A2.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB004776.
XX
PR 13-DEC-1999; 99GB-00029464.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-408419/43.
XX
PT A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX
PS Example 4; Page 87; 646pp; English.
XX
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification

```
XX
SQ   Sequence 10 AA;

Alignment Scores:
Pred. No.:              151          Length:        10
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x AAG94132 (1-10)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||
Db          2 SerSerGlySerSerSerGlySerSer 10

RESULT 68
AAG94130
ID   AAG94130 standard; peptide; 10 AA.
XX
AC   AAG94130;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 324.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB004776.
XX
PR   13-DEC-1999;   99GB-00029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides to
PT   proteins encoded by genes of the human genome, useful in an assay for
PT   screening and identifying of one or more novel peptides which are drug
PT   candidates or pro-drugs.
XX
PS   Example 4; Page 86; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands generated
CC   from the human genome. The complementary peptides interact with their
CC   relevant target proteins encoded in the human genome. They can be used as
CC   reagents in drug discovery and as lead ligands to facilitate drug design
CC   and development. The present sequence is a complementary peptide provided
CC   in the specification
XX
SQ   Sequence 10 AA;

Alignment Scores:
Pred. No.:              151          Length:        10
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x AAG94130 (1-10)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||
Db          1 SerSerGlySerSerSerGlySerSer 9

RESULT 69
ABG68616
ID   ABG68616 standard; peptide; 10 AA.
XX
AC   ABG68616;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Mouse corneodesmosin peptide.
XX
KW   Human; secretory leukoprotease inhibitor; SLPI; corneodesmosin; mouse;
KW   stratum corneum chymotryptic enzyme; envoplakin; desmoplakin; pig; SCCE;
KW   desmocollin 1; adhesion protein; protease; protease inhibitor; eczema;
KW   contact dermatitis; lung atopic asthma; post viral asthma; viral warts;
KW   bronchial hyper-reactivity; chronic obstructive pulmonary disease;
KW   Crohn's disease; ulcerative colitis; coeliac disease; peptic ulceration;
KW   impetigo; meningitis; skin melanoma; squamous cell carcinoma; psoriasis;
KW   basal cell carcinoma; cutaneous lymphoma; skin cancer; lung malignancy;
KW   gastrointestinal tract malignancy; acne vulgaris; psoriasis vulgaris;
KW   antipsoriatic; dermatological; antiinflammatory; antiallergic.
XX
OS   Mus sp.
XX
PN   WO200244736-A2.
XX
PD   06-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-GB005303.
XX
PR   30-NOV-2000; 2000GB-00029225.
PR   07-DEC-2000; 2000GB-00029879.
XX
PA   (MOLE-) MOLECULAR SKINCARE LTD.
XX
PI   Tazi-Ahnini R,  Bavik C,  Ward S,  Duff G,  Cork M;
XX
DR   WPI; 2002-557554/59.
XX
PT   Diagnosing a disease or susceptibility to a disease associated with
PT   abnormal cell-cell adhesion between epithelial cells, by detecting
PT   mutation in nucleic acid encoding adhesion protein, protease or protease
PT   inhibitor.
XX
PS   Example A5; Page 101; 257pp; English.
XX
CC   The invention relates to a method for diagnosis of a disease or
CC   susceptibility to a disease associated with abnormal cell-cell adhesion
CC   between epithelial cells, comprising detecting a mutation in a nucleic
CC   acid encoding an adhesion protein, a protease or a protease inhibitor, or
CC   modulated level of adhesion protein, protease or protease inhibitor, or
CC   its fragment polypeptide. The method is useful for diagnosing a disease
CC   such as eczema, contact dermatitis, lung atopic asthma, post viral
CC   asthma, bronchial hyper-reactivity, chronic obstructive pulmonary
CC   disease, Crohn's disease, ulcerative colitis, coeliac disease, peptic
CC   ulceration, impetigo, viral warts, meningitis, skin melanoma, squamous
CC   cell carcinoma, basal cell carcinoma, cutaneous lymphoma, skin cancer,
CC   malignancy of the gastrointestinal tract, malignancy of the lung,
CC   psoriasis, acne vulgaris and psoriasis vulgaris. This sequence represents
CC   a peptide used in the scope of the invention
XX
SQ   Sequence 10 AA;

Alignment Scores:
Pred. No.:              151          Length:        10
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     5            Gaps:          0

US-10-677-865A-2 (1-1590) x ABG68616 (1-10)
```

```
Qy        142 ATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||
Db          1 IleThrSerProAsnAspProCysLeu 9

RESULT 70
AAR27243
ID   AAR27243 standard; peptide; 13 AA.
XX
AC   AAR27243;
XX
DT   25-MAR-2003  (revised)
DT   25-FEB-1993  (first entry)
XX
DE   Sequence of serine-rich peptide linker.
XX
KW   Fusion protein; serine-rich peptide linker; protease resistant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          3. .7
FT                   /label= repeat unit
FT                   /note= "may be repeated multiple times"
XX
PN   WO9215682-A1.
XX
PD   17-SEP-1992.
XX
PF   27-FEB-1992;   92WO-US001478.
XX
PR   27-FEB-1991;   91US-00662226.
XX
PA   (CREA-) CREATIVE BIOMOLECULES INC.
XX
PI   Huston JS,  Oppermann H,  Timasheff SN;
XX
DR   WPI; 1992-331728/40.
XX
PT   Fusion proteins comprising serine-rich peptide linkers - have improved
PT   solubility in physiological media, resistance to proteolysis and enhanced
PT   refolding properties.
XX
PS   Disclosure; Page 4 and 29; 49pp; English.
XX
CC   AAR27243 is a prefd. linker of the invention. The advantage of such
CC   linkers is that they can improve the refolding properties of the fusion
CC   protein expressed in procaryotes. The linker itself is resistant to
CC   proteolysis. In particular it is used to fuse domains mimicking Vh and Vl
CC   from monoclonal antibody, to produce single chain binding site proteins
CC   (sFv) which dissolve in physiological media, retain their activity at
CC   high concentrations, and resist lysis by endogenous protesases. (Updated
CC   on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 13 AA;

Alignment Scores:
Pred. No.:              146          Length:        13
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x AAR27243 (1-13)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          1 SerGlySerSerSerSerGlySerSer 9

RESULT 71
ADV13591
ID   ADV13591 standard; peptide; 13 AA.
XX
AC   ADV13591;
XX
DT   10-MAR-2005  (first entry)
XX
DE   Human phosphorylated peptide from phophoprotein #1529.
XX
KW   Alzheimers disease; neuroprotective; nootropic; degeneration; tumor;
KW   neoplasm; neurological disease; phosphorylation; protein sequencing;
KW   phosphoprotein.
XX
OS   Homo sapiens.
XX
PN   WO2004108948-A2.
XX
PD   16-DEC-2004.
XX
PF   04-JUN-2004; 2004WO-US017613.
XX
PR   04-JUN-2003; 2003US-0476010P.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Gygi SP;
XX
DR   WPI; 2005-031720/03.
XX
PT   Characterizing phosphorylated polypeptides in a sample comprises
PT   digesting the polypeptides with a protease thus generating test peptides,
PT   and collecting a fraction of test peptides that enriched for positively
PT   charged peptides.
XX
PS   Claim 16; Page 88; 123pp; English.
XX
CC   The invention relates to characterizing phosphorylated polypeptides in a
CC   sample comprising digesting the polypeptides with a protease thus
CC   generating test peptides, and collecting a fraction of test peptides that
CC   enriched for positively charged peptides. Also included are a method
CC   (comprising determining the presence, absence or level of one ore more
CC   phosphorylated peptides as identified above in cells having a cell state
CC   and determining the degree of correlation between the presence, absence
CC   or level of phosphorylated polypeptide with the cell state), an isolated
CC   peptide of 5-50 amino acids comprising an amino acid sequence that is a
CC   subsequence of any of the protein sequences given in the specification
CC   (and which comprise a phosphorylation site within the subsequence), an
CC   isolated polypeptide selected from any of the polypeptides listed in the
CC   specification and is modified at a modification site, an isolated peptide
CC   comprising a mass spectral peak signatures. a method for identifying a
CC   treatment that modulates phosphorylation of an amino acid in a target
CC   polypeptide, a method for generating a peptide standard, a pair of
CC   peptide standards comprising the peptide obtained (where the peptide is
CC   phosphorylated and a corresponding peptide comprising an identical amino
CC   acid sequence but which is not phosphorylated), a system (comprising a
CC   computer memory comprising data files storing information relating to the
CC   identifying characteristics of positively charged peptides, and a data
CC   analysis module capable of executing instructions for organizing and/or
CC   searching the data files), a computer program product (comprising data
CC   relating to the identifying characteristics of positively charged
CC   peptides and comprising instructions for organizing and/or searching the
CC   data), and a method for identifying N-terminal peptides in a sample. The
CC   method is useful for characterizing phosphorylated polypeptides in a
CC   sample. The present sequence is a peptide from a human phosphoprotein,
CC   containing a phosphorylation site, identified by the method of the
CC   invention.
XX
SQ   Sequence 13 AA;

Alignment Scores:
Pred. No.:              146          Length:        13
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
```

```
DB:                    9             Gaps:          0

US-10-677-865A-2 (1-1590) x ADV13591 (1-13)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||
Db          1 SerSerSerSerSerSerGlnSerSer 9

RESULT 72
AAR39572
ID   AAR39572 standard; protein; 15 AA.
XX
AC   AAR39572;
XX
DT   25-MAR-2003  (revised)
DT   07-FEB-1994  (first entry)
XX
DE   Sequence of a new linker.
XX
KW   Single chain Fv polypeptide; VH-VL heterodimer; immunoglobulin; Ig;
KW   biosynthetic single polypeptide chain binding site.
XX
OS   Synthetic.
XX
PN   WO9316185-A2.
XX
PD   19-AUG-1993.
XX
PF   05-FEB-1993;   93WO-US001055.
XX
PR   06-FEB-1992;   92US-00831967.
XX
PA   (CREA-) CREATIVE BIOMOLECULES INC.
PA   (CETU ) CETUS ONCOLOGY CORP.
XX
PI   Huston JS,  Houston LL,  Ring DB,  Oppermann H;
XX
DR   WPI; 1993-272889/34.
DR   N-PSDB; AAQ46087.
XX
PT   New single chain Fv polypeptide binding to C-erbB-2 tumour antigen - for
PT   imaging or treating breast or ovarian cancer etc.
XX
PS   Example; Page 70; 87pp; English.
XX
CC   c-erbB-2 refers to a protein antigen expressed on the surface of tumour
CC   cells. such as breast and ovarian tumour cells, which is an approx.
CC   200,000 mol. wt. acidic glycoprotein having an isolelectric pt. of about
CC   5.3 (see AAQ46083, AAR39568). A single chain Fv (sFv) is a covalently
CC   linked VH-VL heterodimer which is expressed from a gene fusion including
CC   VH- and VL- encoding genes connected by a peptide-encoding linker. Such
CC   linker sequences are set forth in AA residues 116-135 in AAR39569, which
CC   includes part of the 16 AA linker sequences in AAR39572. Using AAQ46084
CC   for the 520C9 monoclonal antibody, a single chain polypeptide can be
CC   produced having a binding affinity for a c-erbB-2 related antigen.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 15 AA;

Alignment Scores:
Pred. No.:              144           Length:        15
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x AAR39572 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 73
AAY90828
ID   AAY90828 standard; peptide; 15 AA.
XX
AC   AAY90828;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Linker amino acid sequence SEQ ID NO:31.
XX
KW   Antigen binding site; immunoglobulin; cancer antigen; immunological;
KW   antibody; tumour; human; mucin; cancer; cytostatic; hybridoma;
KW   specific binding assay; affinity purification; drug targeting;
KW   toxin targeting; imaging; genetic; therapeutic.
XX
OS   Synthetic.
XX
PN   US6054561-A.
XX
PD   25-APR-2000.
XX
PF   07-JUN-1995;   95US-00483749.
XX
PR   08-FEB-1984;   84US-00577976.
PR   11-JAN-1985;   85US-00690750.
PR   21-MAR-1986;   86US-00842476.
PR   08-MAY-1988;   88US-00190778.
PR   11-AUG-1994;   94US-00288981.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Ring DB;
XX
DR   WPI; 2000-338508/29.
XX
PT   Monoclonal antibody capable of binding to human breast cancer antigen
PT   useful for affinity purification, drug or toxin targeting, imaging, and
PT   treating cancer.
XX
PS   Disclosure; Col 77; 57pp; English.
XX
CC   The present invention describes a monoclonal antibody (MAb) (I) that
CC   binds to a human breast cancer antigen that is also bound by MAb 454C11
CC   and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also
CC   described is a hybridoma that produces (I). (I) is useful in specific
CC   binding assays, affinity purification, drug or toxin targeting, imaging,
CC   and genetic or immunological therapeutics for various cancers. The
CC   present sequence represents a linker amino acid sequence which is used in
CC   the exemplification of the present invention
XX
SQ   Sequence 15 AA;

Alignment Scores:
Pred. No.:              144           Length:        15
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     3             Gaps:          0

US-10-677-865A-2 (1-1590) x AAY90828 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 74
AAU08688
ID   AAU08688 standard; peptide; 15 AA.
XX
AC   AAU08688;
XX
DT   18-DEC-2001  (first entry)
```

XX
DE Antibody L19-Fibronectin-2 (IL-2) fusion protein amino acid linker.
XX
KW Angiogenesis; extracellular matrix; ED-B domain; fibronectin; lesion;
KW IL-2; doxorubicin; interleukin-12; Interferon-gamma; IL-12; IFN-gamma;
KW Tumour Necrosis Factor alpha; TNF alpha; tissue factor protein; angioma;
KW electrochemical potential; tumour; cancer; rheumatoid arthritis;
KW diabetic retinopathy; age related muscular degeneration; cytostatic;
KW antirheumatic; antiarthritic; antidiabetic; ophthalmological.
XX
OS Synthetic.
XX
PN WO200162298-A2.
XX
PD 30-AUG-2001.
XX
PF 22-FEB-2001; 2001WO-IB000382.
XX
PR 24-FEB-2000; 2000US-0184767P.
PR 21-DEC-2000; 2000US-0257192P.
XX
PA (PHIL-) PHILOGEN SRL.
XX
PI Zardi L, Neri D, Carnemolla B, Nilsson F, Tarli L, Borsi L;
PI Halin C;
XX
DR WPI; 2001-596713/67.
XX
PT Novel conjugate for treating lesions, comprises a specific binding member
PT specific for extra-cellular matrix component present in lesions, and a
PT molecule that exerts biocidal/cytotoxic effect on target cells in
PT lesions.
XX
PS Example 1; Page 30; 88pp; English.
XX
CC The invention relates to a conjugate comprising a molecule that exerts a
CC biocidal or cytotoxic effect on target cells in lesions of pathological
CC angiogenesis and an antibody (L19) directed against an extracellular
CC matrix component (the ED-B domain of fibronectin) which is present in the
CC lesions. Examples of biocidal or cytotoxic molecules include fibronectin-
CC 2 (IL-2), doxorubicin, interleukin-12 (IL-12), Interferon-gamma (IFN-
CC gamma), Tumour Necrosis Factor alpha (TNF alpha) tissue factor protein.
CC These proteins may interact with a membrane-bound receptor on the target
CC cell or perturb the electrochemical potential of the cell membrane. The
CC treatable disorders include tumours, rheumatoid arthritis, diabetic
CC retinopathy, age related muscular degeneration and angiomas. This
CC sequence represents an amino acid linker encoded by a PCR primer used in
CC the construction of an L19-IL-2 fusion protein
XX
SQ Sequence 15 AA;

Alignment Scores:

Pred. No.:	144	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x AAU08688 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 75
AAU04949
ID AAU04949 standard; peptide; 15 AA.
XX
AC AAU04949;
XX
DT 24-OCT-2001 (first entry)
XX
DE Humanised anti-p185 antibody/IL-2 fusion protein, linker peptide #2.
XX
KW Humanised; antibody; p185; 520C9H; interleukin-2; IL-2; immunoconjugate;
KW cancer; tumour; adenocarcinoma; fusion protein.
XX
OS Synthetic.
XX
PN WO200153354-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001919.
XX
PR 20-JAN-2000; 2000US-0177258P.
XX
PA (CHIR) CHIRON CORP.
PA (HAMI-) HAMILTON CIVIC HOSPITALS RES DEV INC.
PA (HAMI-) HAMILTON REGIONAL CANCER CENT.
XX
PI Austin R, Kwok CS, Ring DB;
XX
DR WPI; 2001-451904/48.
XX
PT Novel immunoconjugate useful for inhibiting tumor cell growth in vivo
PT comprises a humanized anti-p185 antibody linked to an Interleukin-2
PT polypeptide.
XX
PS Disclosure; Page 16; 74pp; English.
XX
CC The sequence represents a linker peptide used to make a humanised anti-
CC p185 (520C9) single chain antibody/human interleukin-2 (IL-2) fusion
CC protein. The fusion protein (or immunoconjugate) is used to inhibit the
CC growth of tumours or cancers particularly those characterised by
CC overexpression of p185 e.g. human adenocarcinomas and malignant and/or
CC benign tumours of the breast, renal system, salivary gland,
CC gastrointestinal tract or gastric tumours
XX
SQ Sequence 15 AA;

Alignment Scores:

Pred. No.:	144	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x AAU04949 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

Search completed: March 21, 2006, 22:19:50
Job time : 185.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:27:04 ; Search time 27.3 Seconds
(without alignments)
4867.023 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3572506

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/absa/ABSSWEB_spool/US10677865/runat_21032006_152110_26939/app_query.fasta_1
-DB=Published_Applications_AA_Main -QFMT=fastan -SUFFIX=oligo.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=50 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=absa05h
-USER=US10677865_@CGN_1_1_134_@runat_21032006_152110_26939 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	65.5	529	4	US-10-316-175-15	Sequence 15, Appl
2	346	65.5	529	5	US-10-677-865-1	Sequence 1, Appli
3	224	42.4	486	4	US-10-316-175-12	Sequence 12, Appl
4	224	42.4	486	5	US-10-723-860-4324	Sequence 4324, Ap
5	224	42.4	486	5	US-10-883-805-16	Sequence 16, Appl
6	205	38.8	529	5	US-10-883-805-6	Sequence 6, Appli
7	198	37.5	529	4	US-10-204-884-3	Sequence 3, Appli
8	123	23.3	529	4	US-10-164-230-61	Sequence 61, Appl
9	94	17.8	219	4	US-10-029-386-34046	Sequence 34046, A
10	81	15.3	81	4	US-10-316-175-73	Sequence 73, Appl
11	69	13.1	146	4	US-10-316-175-71	Sequence 71, Appl
12	59	11.2	111	4	US-10-316-175-72	Sequence 72, Appl
13	58	11.0	58	4	US-10-316-175-70	Sequence 70, Appl
14	55	10.4	55	4	US-10-316-175-69	Sequence 69, Appl
15	15	2.8	15	4	US-10-316-175-17	Sequence 17, Appl
16	15	2.8	15	4	US-10-316-175-19	Sequence 19, Appl
17	14	2.7	16	4	US-10-316-175-16	Sequence 16, Appl
18	11	2.1	11	4	US-10-316-175-24	Sequence 24, Appl
19	11	2.1	238	4	US-10-437-963-119398	Sequence 119398,
20	11	2.1	344	3	US-09-740-288A-8	Sequence 8, Appli
21	11	2.1	344	3	US-09-740-288A-20	Sequence 20, Appl
22	11	2.1	344	4	US-10-431-544-8	Sequence 8, Appli
23	11	2.1	344	4	US-10-431-544-20	Sequence 20, Appl
24	11	2.1	377	3	US-09-740-288A-22	Sequence 22, Appl
25	11	2.1	377	3	US-09-740-288A-24	Sequence 24, Appl
26	11	2.1	377	4	US-10-431-544-22	Sequence 22, Appl
27	11	2.1	377	4	US-10-431-544-24	Sequence 24, Appl
28	11	2.1	377	4	US-10-425-115-256650	Sequence 256650,
29	11	2.1	381	4	US-10-425-114-65547	Sequence 65547, A
30	11	2.1	631	4	US-10-408-765A-1934	Sequence 1934, Ap
31	11	2.1	709	4	US-10-161-493-76	Sequence 76, Appl
32	11	2.1	828	6	US-11-097-143-31512	Sequence 31512, A
33	10	1.9	100	4	US-10-437-963-126164	Sequence 126164,
34	10	1.9	309	4	US-10-369-493-2419	Sequence 2419, Ap
35	10	1.9	344	4	US-10-369-493-2420	Sequence 2420, Ap
36	10	1.9	406	4	US-10-104-047-3291	Sequence 3291, Ap
37	10	1.9	593	3	US-09-964-899-15	Sequence 15, Appl
38	10	1.9	593	5	US-10-975-523-15	Sequence 15, Appl
39	10	1.9	609	4	US-10-408-765A-2321	Sequence 2321, Ap
40	10	1.9	645	3	US-09-964-899-17	Sequence 17, Appl
41	10	1.9	645	5	US-10-975-523-17	Sequence 17, Appl
42	10	1.9	646	3	US-09-964-899-53	Sequence 53, Appl
43	10	1.9	646	5	US-10-975-523-53	Sequence 53, Appl
44	10	1.9	866	4	US-10-081-872-204	Sequence 204, App
45	10	1.9	866	4	US-10-385-305-204	Sequence 204, App
46	10	1.9	5029	5	US-10-732-923-8219	Sequence 8219, Ap
47	10	1.9	5533	6	US-11-097-143-24108	Sequence 24108, A
48	10	1.9	5560	4	US-10-263-929-142	Sequence 142, App
49	10	1.9	5560	6	US-11-097-143-40272	Sequence 40272, A
50	9	1.7	10	3	US-09-572-404B-324	Sequence 324, App

ALIGNMENTS

RESULT 1
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-15

Alignment Scores:
Pred. No.: 6.63e-297 Length: 529
Score: 346.00 Matches: 346
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

```
Query Match:            65.5%       Indels:        0
DB:                     4           Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-316-175-15 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529

RESULT 2
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
;  APPLICANT: SERRE, Guy Bruno Rene
;  APPLICANT:  SIMON, Michel
;  APPLICANT:  WEBER-VIVAT, Marina
;  TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
;  TITLE OF INVENTION:  AND USE THEREOF
;  FILE REFERENCE: 016800-336
;  CURRENT APPLICATION NUMBER: US/10/677,865
;  CURRENT FILING DATE:  2003-10-03
;  PRIOR APPLICATION NUMBER: US/09/381,656
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: FR 97/03899
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-10-677-865-1

Alignment Scores:
Pred. No.:              6.63e-297       Length:         529
Score:                  346.00          Matches:        346
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            65.5%           Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-677-865-1 (1-529)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343
```

```
Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy       1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy       1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 SerProProGlyThrGlySerPheSerSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy       1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy       1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy       1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy       1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy       1570 GAGTTACTCGACAGTCCA 1587
             ||||||||||||||||||
Db        524 GluLeuLeuAspSerPro 529
```

RESULT 3
US-10-316-175-12
; Sequence 12, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-12

Alignment Scores:

Pred. No.:	6.14e-189	Length:	486
Score:	224.00	Matches:	224
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	42.4%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-316-175-12 (1-486)

```
Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 187

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 207

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 227

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 247

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 267

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 287

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 307

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 327

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 347

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 367

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 387

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        388 IleSerSerSer 391
```

RESULT 4
US-10-723-860-4324
; Sequence 4324, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4324
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4324

```
Alignment Scores:
Pred. No.:              6.14e-189       Length:         486
Score:                  224.00          Matches:        224
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            42.4%           Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-723-860-4324 (1-486)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 187

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 207

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 227

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 247

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 267

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 287

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 307

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 327

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 347

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 367

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 387

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        388 IleSerSerSer 391

RESULT 5
US-10-883-805-16
; Sequence 16, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 16
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No: g414810
US-10-883-805-16

Alignment Scores:
Pred. No.:              6.14e-189       Length:         486
Score:                  224.00          Matches:        224
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            42.4%           Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-883-805-16 (1-486)

Qy        550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 187

Qy        610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 207

Qy        670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 227

Qy        730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 247

Qy        790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 267

Qy        850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 287

Qy        910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 307

Qy        970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 327

Qy       1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 347

Qy       1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 367

Qy       1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        368 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 387

Qy       1210 ATTTCCAGCAGC 1221
              ||||||||||||
Db        388 IleSerSerSer 391

RESULT 6
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

Alignment Scores:
Pred. No.:              3.95e-172      Length:          529
Score:                  205.00         Matches:         205
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            38.8%          Indels:          0
DB:                     5              Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-883-805-6 (1-529)

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIle 222

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSer 242

Qy        727 TCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 SerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySer 262

Qy        787 GGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysPro 282

Qy        847 TGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySer 302

Qy        907 TCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGC 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 SerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGly 322

Qy        967 TACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGly 342

Qy       1027 CCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCA 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 ProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAla 362

Qy       1087 GCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGly 382

Qy       1147 GGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCT 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSer 402

Qy       1207 AGCATTTCCAGCAGC 1221
              ||||||||||||||
Db        403 SerIleSerSerSer 407

RESULT 7
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (143)..(143)
;   OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (401)..(401)
;   OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (409)..(409)
;   OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
```

```
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (410)..(410)
;   OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (527)..(527)
;   OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

Alignment Scores:
Pred. No.:              6.18e-166       Length:         529
Score:                  198.00          Matches:        198
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            37.5%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-204-884-3 (1-529)

Qy        607 GGGGTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATC 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GlyValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIle 222

Qy        667 CCCGACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 ProAspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSer 242

Qy        727 TCCCACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 SerHisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySer 262

Qy        787 GGTGCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCC 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GlyAlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysPro 282

Qy        847 TGTCCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 CysProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySer 302

Qy        907 TCTGACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGC 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 SerAspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGly 322

Qy        967 TACTTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TyrPheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGly 342

Qy       1027 CCCCCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCA 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 ProProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAla 362

Qy       1087 GCTTCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AlaSerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGly 382

Qy       1147 GGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GlySerThrGlySerLysGlyProCysSerProSerSerSerArgValProSer 400

RESULT 8
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

Alignment Scores:
Pred. No.:              1.45e-99        Length:         529
Score:                  123.00          Matches:        123
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            23.3%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-164-230-61 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy         181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy         241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGly 100

Qy         301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy         361 TCTGGCTCT 369
                |||||||||
Db         121 SerGlySer 123

RESULT 9
US-10-029-386-34046
; Sequence 34046, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34046
;   LENGTH: 219
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000510.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.21
;   OTHER INFORMATION: SWISSPROT HIT: Q15517, EVALUE 1.00e-24
US-10-029-386-34046

Alignment Scores:
Pred. No.:              7.53e-74        Length:         219
Score:                  94.00           Matches:        94
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            17.8%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-029-386-34046 (1-219)

Qy          88 ACCTTGGCTAAGAGCATTGGCACCTTCTCAGACCCCTGTAAGGACCCCACGCGTATCACC 147
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ThrLeuAlaLysSerIleGlyThrPheSerAspProCysLysAspProThrArgIleThr 20

Qy         148 TCCCCTAACGACCCCTGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGT 207
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 SerProAsnAspProCysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSer 40

Qy         208 GGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGT 267
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 GlySerSerSerSerGlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSer 60

Qy         268 GGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGA 327
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGly 80

Qy         328 ACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCT 369
                |||||||||||||||||||||||||||||||||||||||||
Db          81 ThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySer 94

RESULT 10
US-10-316-175-73
; Sequence 73, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-73

Alignment Scores:
Pred. No.:              2.76e-62        Length:         81
Score:                  81.00           Matches:        81
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.3%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-73 (1-81)

Qy        1123 GGTGGGGTCCAGCTCTGTGGAGGCGGCTCCACGGGCTCCAAGGGACCCTGCTCTCCCTCC 1182
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GlyGlyValGlnLeuCysGlyGlyGlySerThrGlySerLysGlyProCysSerProSer 20

Qy        1183 AGTTCTCGAGTCCCCAGCAGTTCTAGCATTTCCAGCAGCGCCGGTTCACCCTACCATCCC 1242
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 SerSerArgValProSerSerSerSerIleSerSerSerAlaGlySerProTyrHisPro 40

Qy        1243 TGCGGCAGTGCTTCCCAGAGCCCCTGCTCCCCACCAGGCACCGGCTCCTTCAGCAGCAGC 1302
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 CysGlySerAlaSerGlnSerProCysSerProProGlyThrGlySerPheSerSerSer 60

Qy        1303 TCCAGTTCCCAATCGAGTGGCAAAATCATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCT 1362
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 SerSerSerGlnSerSerGlyLysIleIleLeuGlnProCysGlySerLysSerSerSer 80

Qy        1363 TCT 1365
                |||
Db          81 Ser 81

RESULT 11
US-10-316-175-71
; Sequence 71, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
```

```
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
;   LENGTH: 146
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-71

Alignment Scores:
Pred. No.:              1.06e-51       Length:         146
Score:                  69.00          Matches:        69
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            13.1%          Indels:         0
DB:                     4              Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-71 (1-146)

Qy        163 TGCCTCACTGGGAAGGGTGACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCT 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CysLeuThrGlyLysGlyAspSerSerGlyPheSerSerTyrSerGlySerSerSerSer 20

Qy        223 GGCAGCTCCATTTCCAGTGCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGA 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlySerSerIleSerSerAlaArgSerSerGlyGlyGlySerSerGlySerSerSerGly 40

Qy        283 TCCAGCATTGCCCAGGGTGGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAG 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerSerIleAlaGlnGlyGlySerAlaGlySerPheLysProGlyThrGlyTyrSerGln 60

Qy        343 GTCAGCTACTCCTCCGGATCTGGCTCT 369
              |||||||||||||||||||||||||||
Db         61 ValSerTyrSerSerGlySerGlySer 69

RESULT 12
US-10-316-175-72
; Sequence 72, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72
;   LENGTH: 111
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-72

Alignment Scores:
Pred. No.:              7.77e-43       Length:         111
Score:                  59.00          Matches:        59
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            11.2%          Indels:         0
DB:                     4              Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-72 (1-111)

Qy        193 TTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGTGCCAGAAGCTCT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSerAlaArgSerSer 20

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGTGGTTCTGCAGGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyGlySerSerGlySerSerSerGlySerSerIleAlaGlnGlyGlySerAlaGly 40

Qy        313 TCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGATCTGGCTCT 369
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGlySerGlySer 59

RESULT 13
US-10-316-175-70
; Sequence 70, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-70

Alignment Scores:
Pred. No.:              6.53e-42       Length:         58
Score:                  58.00          Matches:        58
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            11.0%          Indels:         0
DB:                     4              Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-70 (1-58)

Qy       1414 GATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCCAAGCCCTGTGGCTCCAGCAGTGCT 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspGlySerProHisProAspProSerAlaGlyAlaLysProCysGlySerSerSerAla 20

Qy       1474 GGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTAGCCCAAGTGAAGCCTCTGGGGCCC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyLysIleProCysArgSerIleArgAspIleLeuAlaGlnValLysProLeuGlyPro 40

Qy       1534 CAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGAGAGTTACTCGACAGTCCA 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlnLeuAlaAspProGluValPheLeuProGlnGlyGluLeuLeuAspSerPro 58

RESULT 14
US-10-316-175-69
; Sequence 69, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
;   LENGTH: 55
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-69

Alignment Scores:
Pred. No.:              2.97e-39        Length:         55
Score:                  55.00           Matches:        55
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            10.4%           Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-69 (1-55)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCCTGC 165
              |||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCys 55

RESULT 15
US-10-316-175-17
; Sequence 17, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-17

Alignment Scores:
Pred. No.:              0.000888        Length:         15
Score:                  15.00           Matches:        15
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.8%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-17 (1-15)

Qy        196 AGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 15

RESULT 16
US-10-316-175-19
; Sequence 19, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-19

Alignment Scores:
Pred. No.:              0.000888        Length:         15
Score:                  15.00           Matches:        15
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.8%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-19 (1-15)

Qy        436 TCTCACTCGGGAAGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerHisSerGlySerSerSerSerHisSerSerSerSerSerSer 15

RESULT 17
US-10-316-175-16
; Sequence 16, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-16

Alignment Scores:
Pred. No.:              0.00675         Length:         16
Score:                  14.00           Matches:        14
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            2.7%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-316-175-16 (1-16)

Qy         55 GCACTGCTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCT 96
              ||||||||||||||||||||||||||||||||||||||||||
Db          3 AlaLeuLeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAla 16
```

```
RESULT 18
US-10-316-175-24
; Sequence 24, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-24

Alignment Scores:
Pred. No.:              3.21          Length:          11
Score:                  11.00         Matches:         11
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            2.1%          Indels:          0
DB:                     4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-316-175-24 (1-11)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              |||||||||||||||||||||||||||||||||
Db          1 SerPheSerSerSerSerSerSerGlnSerSer 11

RESULT 19
US-10-437-963-119398
; Sequence 119398, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 119398
;   LENGTH: 238
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_22619C.1.pep
US-10-437-963-119398

Alignment Scores:
Pred. No.:              2.09          Length:          238
Score:                  11.00         Matches:         11
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            2.1%          Indels:          0
DB:                     4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-437-963-119398 (1-238)

Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||||||||
Db        155 SerSerSerSerHisSerSerSerSerSerSer 165

RESULT 20
US-09-740-288A-8
; Sequence 8, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  APPLICANT:  Kinney, Anthony
;  APPLICANT:  Miao, Guo-Hua
;  APPLICANT:  Orozco, Emil
;  TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
;  FILE REFERENCE: BB1429 US NA
;  CURRENT APPLICATION NUMBER: US/09/740,288A
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER: US 60/172929
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-740-288A-8

Alignment Scores:
Pred. No.:              1.99          Length:          344
Score:                  11.00         Matches:         11
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            2.1%          Indels:          0
DB:                     3             Gaps:            0

US-10-677-865A-2 (1-1590) x US-09-740-288A-8 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 21
US-09-740-288A-20
; Sequence 20, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  APPLICANT:  Kinney, Anthony
;  APPLICANT:  Miao, Guo-Hua
;  APPLICANT:  Orozco, Emil
;  TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
;  FILE REFERENCE: BB1429 US NA
;  CURRENT APPLICATION NUMBER: US/09/740,288A
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER: US 60/172929
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-740-288A-20

Alignment Scores:
Pred. No.:              1.99          Length:          344
Score:                  11.00         Matches:         11
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            2.1%          Indels:          0
```

DB: 3 Gaps: 0

US-10-677-865A-2 (1-1590) x US-09-740-288A-20 (1-344)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 22
US-10-431-544-8
; Sequence 8, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Gou-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Zea mays
US-10-431-544-8

Alignment Scores:

Pred. No.:	1.99	Length:	344
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-431-544-8 (1-344)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 23
US-10-431-544-20
; Sequence 20, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Gou-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Zea mays
US-10-431-544-20

Alignment Scores:

Pred. No.:	1.99	Length:	344
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-431-544-20 (1-344)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 24
US-09-740-288A-22
; Sequence 22, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-09-740-288A-22

Alignment Scores:

Pred. No.:	1.96	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-740-288A-22 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 25
US-09-740-288A-24
; Sequence 24, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-09-740-288A-24

Alignment Scores:

Pred. No.:	1.96	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2.1% Indels: 0
DB: 3 Gaps: 0

US-10-677-865A-2 (1-1590) x US-09-740-288A-24 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 26
US-10-431-544-22
; Sequence 22, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Gou-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-10-431-544-22

Alignment Scores:
Pred. No.: 1.96 Length: 377
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2.1% Indels: 0
DB: 4 Gaps: 0

US-10-677-865A-2 (1-1590) x US-10-431-544-22 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 27
US-10-431-544-24
; Sequence 24, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Gou-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-10-431-544-24

Alignment Scores:
Pred. No.: 1.96 Length: 377
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2.1% Indels: 0
DB: 4 Gaps: 0

US-10-677-865A-2 (1-1590) x US-10-431-544-24 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 28
US-10-425-115-256650
; Sequence 256650, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 256650
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165654C.1.pep
US-10-425-115-256650

Alignment Scores:
Pred. No.: 1.96 Length: 377
Score: 11.00 Matches: 11
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2.1% Indels: 0
DB: 4 Gaps: 0

US-10-677-865A-2 (1-1590) x US-10-425-115-256650 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 29
US-10-425-114-65547
; Sequence 65547, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65547
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4767-018-H9_FLI.pep
US-10-425-114-65547

```
Alignment Scores:
Pred. No.:                1.96          Length:          381
Score:                    11.00         Matches:         11
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              2.1%          Indels:          0
DB:                       4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-425-114-65547 (1-381)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         23 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 33

RESULT 30
US-10-408-765A-1934
; Sequence 1934, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
;  APPLICANT: Ghosh, Soumitra S.
;  APPLICANT:  Fahy, Eoin D.
;  APPLICANT:  Zhang, Bing
;  APPLICANT:  Gibson, Bradford W.
;  APPLICANT:  Taylor, Steven W.
;  APPLICANT:  Glenn, Gary M.
;  APPLICANT:  Warnock, Dale E.
;  TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;  TITLE OF INVENTION:  IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;  FILE REFERENCE: 660088.465
;  CURRENT APPLICATION NUMBER: US/10/408,765A
;  CURRENT FILING DATE:  2003-04-04
;  NUMBER OF SEQ ID NOS: 3077
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1934
;   LENGTH: 631
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-408-765A-1934

Alignment Scores:
Pred. No.:                1.83          Length:          631
Score:                    11.00         Matches:         11
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              2.1%          Indels:          0
DB:                       4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-408-765A-1934 (1-631)

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||
Db         92 SerSerGlyGlyGlySerSerGlySerSerSer 102

RESULT 31
US-10-161-493-76
; Sequence 76, Application US/10161493
; Publication No. US20040018555A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Li, Li
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Pena, Carol EA
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Leite, Mario W
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Taupier Jr., Raymond J
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Voss, Edward Z
;  APPLICANT:  Boldog, Ferenc L
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  TITLE OF INVENTION: No. US20040018555A1el Antibodies that Bind to Antigenic Polypept
;  TITLE OF INVENTION:  Acids Encoding the Antigens, and Methods of Use
;  FILE REFERENCE: 21402-377A
;  CURRENT APPLICATION NUMBER: US/10/161,493
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/337,524
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,575
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: 60/359,151
;  PRIOR FILING DATE: 2002-02-21
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,573
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/341,143
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 299
; SEQ ID NO 76
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-493-76

Alignment Scores:
Pred. No.:                1.8           Length:          709
Score:                    11.00         Matches:         11
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              2.1%          Indels:          0
DB:                       4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-161-493-76 (1-709)

Qy        247 AGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              |||||||||||||||||||||||||||||||||
Db        170 SerSerGlyGlyGlySerSerGlySerSerSer 180

RESULT 32
US-11-097-143-31512
; Sequence 31512, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
```

```
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31512
;   LENGTH: 828
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-31512

Alignment Scores:
Pred. No.:                1.76          Length:          828
Score:                    11.00         Matches:         11
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              2.1%          Indels:          0
DB:                       6             Gaps:            0

US-10-677-865A-2 (1-1590) x US-11-097-143-31512 (1-828)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||||
Db        301 GlyGlySerSerGlySerSerSerGlySerSer 311

RESULT 33
US-10-437-963-126164
; Sequence 126164, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126164
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(100)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_28738C.1.pep
US-10-437-963-126164

Alignment Scores:
Pred. No.:                18.1          Length:          100
Score:                    10.00         Matches:         10
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.9%          Indels:          0
DB:                       4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-437-963-126164 (1-100)

Qy        452 GCAGCTCTCATTCGAGCAGCAGCAGCAGCT 481
              ||||||||||||||||||||||||||||||
Db         44 AlaAlaLeuIleArgAlaAlaAlaAlaAla 53

RESULT 34
US-10-369-493-2419
; Sequence 2419, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2419
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2419

Alignment Scores:
Pred. No.:                15.5          Length:          309
Score:                    10.00         Matches:         10
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.9%          Indels:          0
DB:                       4             Gaps:            0

US-10-677-865A-2 (1-1590) x US-10-369-493-2419 (1-309)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        150 SerSerSerHisSerSerSerSerSerSer 159

RESULT 35
US-10-369-493-2420
; Sequence 2420, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2420
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2420
```

```
Alignment Scores:
Pred. No.:              15.3          Length:         344
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-369-493-2420 (1-344)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        185 SerSerSerHisSerSerSerSerSerSer 194

RESULT 36
US-10-104-047-3291
; Sequence 3291, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20030236392A1el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3291
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-3291

Alignment Scores:
Pred. No.:              14.9          Length:         406
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-104-047-3291 (1-406)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81

RESULT 37
US-09-964-899-15
; Sequence 15, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Dalia et al.
;  TITLE OF INVENTION: Identification of Genes Involved in
;  TITLE OF INVENTION:  Alzheimer's Disease Using Drosophila Melanogaster
;  FILE REFERENCE: 4-31612 A
;  CURRENT APPLICATION NUMBER: US/09/964,899
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: 60/236,893
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/298,309
;  PRIOR FILING DATE: 2001-06-14
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 593
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-964-899-15

Alignment Scores:
Pred. No.:              14.1          Length:         593
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     3             Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-964-899-15 (1-593)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        261 GlyGlySerSerGlySerSerSerGlySer 270

RESULT 38
US-10-975-523-15
; Sequence 15, Application US/10975523
; Publication No. US20050138676A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Dalia et al.
;  TITLE OF INVENTION: Identification of Genes Involved in
;  TITLE OF INVENTION:  Alzheimer's Disease Using Drosophila Melanogaster
;  FILE REFERENCE: 4-31612 A
;  CURRENT APPLICATION NUMBER: US/10/975,523
;  CURRENT FILING DATE:  2004-10-28
;  PRIOR APPLICATION NUMBER: US/09/964,899
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/236,893
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/298,309
;  PRIOR FILING DATE: 2001-06-14
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 593
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-975-523-15

Alignment Scores:
Pred. No.:              14.1          Length:         593
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     5             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-975-523-15 (1-593)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        261 GlyGlySerSerGlySerSerSerGlySer 270

RESULT 39
US-10-408-765A-2321
; Sequence 2321, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
;  APPLICANT: Ghosh, Soumitra S.
;  APPLICANT:  Fahy, Eoin D.
;  APPLICANT:  Zhang, Bing
;  APPLICANT:  Gibson, Bradford W.
;  APPLICANT:  Taylor, Steven W.
;  APPLICANT:  Glenn, Gary M.
;  APPLICANT:  Warnock, Dale E.
;  TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;  TITLE OF INVENTION:  IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;  FILE REFERENCE: 660088.465
;  CURRENT APPLICATION NUMBER: US/10/408,765A
;  CURRENT FILING DATE:  2003-04-04
;  NUMBER OF SEQ ID NOS: 3077
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2321
;   LENGTH: 609
;   TYPE: PRT
```

; ORGANISM: Homo sapiens
US-10-408-765A-2321

Alignment Scores:

Pred. No.:	14.1	Length:	609
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-408-765A-2321 (1-609)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        275 GlyGlySerSerGlySerSerSerGlySer 284
```

RESULT 40
US-09-964-899-17
; Sequence 17, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-17

Alignment Scores:

Pred. No.:	14	Length:	645
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-964-899-17 (1-645)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        311 GlyGlySerSerGlySerSerSerGlySer 320
```

RESULT 41
US-10-975-523-17
; Sequence 17, Application US/10975523
; Publication No. US20050138676A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/10/975,523
; CURRENT FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: US/09/964,899
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-975-523-17

Alignment Scores:

Pred. No.:	14	Length:	645
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-975-523-17 (1-645)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        311 GlyGlySerSerGlySerSerSerGlySer 320
```

RESULT 42
US-09-964-899-53
; Sequence 53, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-964-899-53

Alignment Scores:

Pred. No.:	14	Length:	646
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-964-899-53 (1-646)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        312 GlyGlySerSerGlySerSerSerGlySer 321
```

RESULT 43
US-10-975-523-53
; Sequence 53, Application US/10975523
; Publication No. US20050138676A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/10/975,523
; CURRENT FILING DATE: 2004-10-28
; PRIOR APPLICATION NUMBER: US/09/964,899
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309

```
;  PRIOR FILING DATE: 2001-06-14
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-975-523-53

Alignment Scores:
Pred. No.:              14              Length:         646
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-975-523-53 (1-646)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        312 GlyGlySerSerGlySerSerSerGlySer 321

RESULT 44
US-10-081-872-204
; Sequence 204, Application US/10081872
; Publication No. US20030125534A1
; GENERAL INFORMATION:
;  APPLICANT: Callen, Walter
;  APPLICANT:  Richardson, Toby
;  APPLICANT:  Frey, Gerhard
;  APPLICANT:  Short, Jay M.
;  APPLICANT:  Mathur, Eric J.
;  APPLICANT:  Gray, Kevin A.
;  APPLICANT:  Kerovuo, Janne S.
;  APPLICANT:  Slupska, Malgorzata
;  TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
;  TITLE OF INVENTION:  AND METHODS OF USE THEREOF
;  FILE REFERENCE: 09010-108001
;  CURRENT APPLICATION NUMBER: US/10/081,872
;  CURRENT FILING DATE:  2002-02-21
;  PRIOR APPLICATION NUMBER: US 60/270,495
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: US 60/270,496
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: US 60/291,122
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 321
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204
;   LENGTH: 866
;   TYPE: PRT
;   ORGANISM: Environmental
US-10-081-872-204

Alignment Scores:
Pred. No.:              13.4            Length:         866
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-081-872-204 (1-866)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        462 GlySerSerGlySerSerSerGlySerSer 471

RESULT 45
US-10-385-305-204
; Sequence 204, Application US/10385305
; Publication No. US20040018607A1
; GENERAL INFORMATION:
;  APPLICANT: Callen, Walter
;  APPLICANT:  Richardson, Toby
;  APPLICANT:  Frey, Gerhard
;  APPLICANT:  Short, Jay M.
;  APPLICANT:  Mathur, Eric J.
;  APPLICANT:  Gray, Kevin A.
;  APPLICANT:  Kerovuo, Janne S.
;  APPLICANT:  Slupska, Malgorzata
;  TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
;  TITLE OF INVENTION:  AND METHODS OF USE THEREOF
;  FILE REFERENCE: 09010-108001
;  CURRENT APPLICATION NUMBER: US/10/385,305
;  CURRENT FILING DATE:  2003-03-06
;  PRIOR APPLICATION NUMBER: US/10/081,872
;  PRIOR FILING DATE: 2002-02-21
;  PRIOR APPLICATION NUMBER: US 60/270,495
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: US 60/270,496
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: US 60/291,122
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 321
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204
;   LENGTH: 866
;   TYPE: PRT
;   ORGANISM: Environmental
US-10-385-305-204

Alignment Scores:
Pred. No.:              13.4            Length:         866
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-385-305-204 (1-866)

Qy        259 GGCTCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||||
Db        462 GlySerSerGlySerSerSerGlySerSer 471

RESULT 46
US-10-732-923-8219
; Sequence 8219, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;  APPLICANT: Edgerton, Michael D
;  TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;  FILE REFERENCE: 38-15(52796)C
;  CURRENT APPLICATION NUMBER: US/10/732,923
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: 10/310,154
;  PRIOR FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8219
;   LENGTH: 5029
;   TYPE: PRT
;   ORGANISM: Anopheles gambiae str. PEST
US-10-732-923-8219

Alignment Scores:
Pred. No.:              10.5            Length:         5029
Score:                  10.00           Matches:        10
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.9%            Indels:         0
DB:                     5               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-732-923-8219 (1-5029)
```

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db       4815 SerSerSerHisSerSerSerSerSerSer 4824

RESULT 47
US-11-097-143-24108
; Sequence 24108, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24108
;   LENGTH: 5533
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-24108

Alignment Scores:
Pred. No.:              10.4          Length:         5533
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     6             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-097-143-24108 (1-5533)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 48
US-10-263-929-142
; Sequence 142, Application US/10263929
; Publication No. US20040067535A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Jaeseob
;  APPLICANT:  Galant, Ron
;  TITLE OF INVENTION: Alzheimer's Disease Linked Genes
;  FILE REFERENCE: LSD-07417
;  CURRENT APPLICATION NUMBER: US/10/263,929
;  CURRENT FILING DATE:  2002-10-03
;  NUMBER OF SEQ ID NOS: 213
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 142
;   LENGTH: 5560
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-263-929-142

Alignment Scores:
Pred. No.:              10.4          Length:         5560
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-263-929-142 (1-5560)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 49
US-11-097-143-40272
; Sequence 40272, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40272
;   LENGTH: 5560
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-40272

Alignment Scores:
Pred. No.:              10.4          Length:         5560
Score:                  10.00         Matches:        10
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.9%          Indels:         0
DB:                     6             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-097-143-40272 (1-5560)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||
Db        456 SerSerSerHisSerSerSerSerSerSer 465

RESULT 50
US-09-572-404B-324
; Sequence 324, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
```

```
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 324
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in SOX11 at 346-355 and may interact with Sequen
;   OTHER INFORMATION: in this patent.
US-09-572-404B-324

Alignment Scores:
Pred. No.:              192          Length:        10
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     3            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-572-404B-324 (1-10)

Qy       262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
             ||||||||||||||||||||||||||||
Db         1 SerSerGlySerSerSerGlySerSer 9

RESULT 51
US-09-572-404B-326
; Sequence 326, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 326
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in SOX11 at 345-354 and may interact with Sequen
;   OTHER INFORMATION: in this patent.
US-09-572-404B-326

Alignment Scores:
Pred. No.:              192          Length:        10
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     3            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-572-404B-326 (1-10)

Qy       262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
             ||||||||||||||||||||||||||||
Db         2 SerSerGlySerSerSerGlySerSer 10

RESULT 52
US-10-862-195-1580
; Sequence 1580, Application US/10862195
; Publication No. US20050164324A1
; GENERAL INFORMATION:
;  APPLICANT: GYGI, STEVEN P.
;  TITLE OF INVENTION: SYSTEMS, METHODS AND KITS FOR CHARACTERIZING PHOSPHOPROTEOMES
;  FILE REFERENCE: 58890(70207)
;  CURRENT APPLICATION NUMBER: US/10/862,195
;  CURRENT FILING DATE:  2004-06-04
;  PRIOR APPLICATION NUMBER: 60/476,010
;  PRIOR FILING DATE: 2003-06-04
;  NUMBER OF SEQ ID NOS: 2245
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1580
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: See specification as filed for preferred embodiments
;   OTHER INFORMATION: and description of phosphorylation sites
US-10-862-195-1580

Alignment Scores:
Pred. No.:              185          Length:        13
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     5            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-862-195-1580 (1-13)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||||
Db          1 SerSerSerSerSerSerGlnSerSer 9

RESULT 53
US-09-144-886-3
; Sequence 3, Application US/09144886
; Patent No. US20020155114A1
; GENERAL INFORMATION:
;  APPLICANT: Marks, James D
;  APPLICANT: Amersdorfer, Peter
;  TITLE OF INVENTION: Therapeutic Monoclonal Antibodies That Neutralize
;  TITLE OF INVENTION: Botulinum Neurotoxins
;  FILE REFERENCE: 2500.117USO
;  CURRENT APPLICATION NUMBER: US/09/144,886
;  CURRENT FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: [(Ser)4Gly]3
;   OTHER INFORMATION: linker
US-09-144-886-3

Alignment Scores:
Pred. No.:              181          Length:        15
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     3            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-144-886-3 (1-15)

Qy       205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
             ||||||||||||||||||||||||||||
Db         4 SerGlySerSerSerSerGlySerSer 12

RESULT 54
US-09-887-853-8
; Sequence 8, Application US/09887853
; Patent No. US20020168375A1
;   GENERAL INFORMATION:
;        APPLICANT: Huston, James S.
;                   Oppermann, Hermann
;                   Houston, L. L.
;                   Ring, David B.
```

```
;       TITLE OF INVENTION: Biosynthetic Binding Proteins For
;                           Imaging
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Testa, Hurwitz & Thibeault/Patent Department
;            STREET: Exchange Place, 53 State Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/887,853
;            FILING DATE: 21-Jun-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/133,804
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kelley, Robin D.
;            REGISTRATION NUMBER: 34,637
;            REFERENCE/DOCKET NUMBER: 2054/22
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 617-248-7477
;            TELEFAX: 617-248-7100
;   INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 15 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       FEATURE:
;            NAME/KEY:  Protein
;            LOCATION:  1..15
;            OTHER INFORMATION:  /note= "LINKER 2"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-887-853-8

Alignment Scores:
Pred. No.:              181          Length:        15
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     3            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-887-853-8 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 55
US-10-204-581-1
; Sequence 1, Application US/10204581
; Publication No. US20040013640A1
; GENERAL INFORMATION:
;  APPLICANT: Zardi, Luciano
;  APPLICANT:  Neri, Dario
;  APPLICANT:  Carnemolla, Barbara
;  APPLICANT:  Nilsson, Fredrik
;  APPLICANT:  Tarli, Lorenzo
;  APPLICANT:  Borsi, Laura
;  APPLICANT:  Halin, Cornelia
;  TITLE OF INVENTION: Compositions and methods for treatment of angiogenesis
;  TITLE OF INVENTION:  in pathological lesions
;  FILE REFERENCE: 0380-P02964US0
;  CURRENT APPLICATION NUMBER: US/10/204,581
;  CURRENT FILING DATE:  2002-12-13
;  PRIOR APPLICATION NUMBER: PCT/IB01/00382
;  PRIOR FILING DATE: 2001-02-22
;  PRIOR APPLICATION NUMBER: US 60/184,767
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/257,192
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Linker
US-10-204-581-1

Alignment Scores:
Pred. No.:              181          Length:        15
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-204-581-1 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 56
US-10-632-706-3
; Sequence 3, Application US/10632706
; Publication No. US20040175385A1
; GENERAL INFORMATION:
;  APPLICANT: MARKS, JAMES D.
;  APPLICANT:  AMERSDORFER, PETER
;  TITLE OF INVENTION: THERAPEUTIC MONOCLONAL ANTIBODIES THAT NEUTRALIZE BOTULINUM
;  TITLE OF INVENTION:  NEUROTOXINS
;  FILE REFERENCE: 407T-895120US
;  CURRENT APPLICATION NUMBER: US/10/632,706
;  CURRENT FILING DATE:  2003-08-01
;  PRIOR APPLICATION NUMBER: US 60/400,721
;  PRIOR FILING DATE: 2002-08-01
;  PRIOR APPLICATION NUMBER: US 09/144,806
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: protein linker
US-10-632-706-3

Alignment Scores:
Pred. No.:              181          Length:        15
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-632-706-3 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 57
US-10-683-547-8
```

; Sequence 8, Application US/10683547
; Publication No. US20050058638A1
; GENERAL INFORMATION:
; APPLICANT: Huston, J.
; APPLICANT: Houston, L.L.
; APPLICANT: Ring, D.
; APPLICANT: Oppermann, H.
; TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING
; FILE REFERENCE: CIBT-P01-130
; CURRENT APPLICATION NUMBER: US/10/683,547
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US/09/558,741
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 07/831,967
; PRIOR FILING DATE: 1992-02-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker 2
US-10-683-547-8

Alignment Scores:

Pred. No.:	181	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-683-547-8 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 58
US-10-818-694-18
; Sequence 18, Application US/10818694
; Publication No. US20050089931A1
; GENERAL INFORMATION:
; APPLICANT: Bamdad, Cynthia C.
; TITLE OF INVENTION: THE USE OF SELF-ASSEMBLED MONOLAYERS TO
; TITLE OF INVENTION: PROBE THE STRUCTURE OF A TARGET MOLECULE
; FILE REFERENCE: M1015/7004/TJO
; CURRENT APPLICATION NUMBER: US/10/818,694
; CURRENT FILING DATE: 2004-04-06
; PRIOR APPLICATION NUMBER: U.S. 60/087,766
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker
US-10-818-694-18

Alignment Scores:

Pred. No.:	181	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	5	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-818-694-18 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 59
US-09-906-206A-68
; Sequence 68, Application US/09906206A
; Publication No. US20050239170A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Urban, Robert G.
; APPLICANT: Aziz, Nazneen
; APPLICANT: Chen, Hongmin
; APPLICANT: Etemad-Moghadam, Bijan
; APPLICANT: Yin, Peng
; TITLE OF INVENTION: ALPHA-MSH RELATED COMPOUNDS AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 08191-015001
; CURRENT APPLICATION NUMBER: US/09/906,206A
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,381
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/226,382
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 60/238,380
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/258,764
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: US 60/298,317
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-906-206A-68

Alignment Scores:

Pred. No.:	178	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	3	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-906-206A-68 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 60
US-10-081-400-4
; Sequence 4, Application US/10081400
; Publication No. US20020155998A1
; GENERAL INFORMATION:
; APPLICANT: Young, Michael
; APPLICANT: Meade, Harry
; APPLICANT: Krane, Ian
; TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
; FILE REFERENCE: 10275/041001
; CURRENT APPLICATION NUMBER: US/10/081,400
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/333,213
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated linker sequence

US-10-081-400-4

Alignment Scores:

Pred. No.:	178	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-081-400-4 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 61
US-10-074-956-19
; Sequence 19, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-19

Alignment Scores:

Pred. No.:	178	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-074-956-19 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 62
US-10-077-213-4
; Sequence 4, Application US/10077213
; Publication No. US20030049694A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Chung-Hsiun
; APPLICANT: Lin, Rong-Hwa
; APPLICANT: Hsu, Pei-Ling
; TITLE OF INVENTION: PRODUCTION OF FUSION PROTEINS AND USE
; TITLE OF INVENTION: FOR IDENTIFYING BINDING MOLECULES
; FILE REFERENCE: 13062-002001
; CURRENT APPLICATION NUMBER: US/10/077,213
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/318,474
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence

US-10-077-213-4

Alignment Scores:

Pred. No.:	178	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-077-213-4 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 63
US-10-077-210-4
; Sequence 4, Application US/10077210
; Publication No. US20030049801A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Chung-Hsiun
; APPLICANT: Lin, Rong-Hwa
; APPLICANT: Hsu, Pei-Ling
; TITLE OF INVENTION: PRODUCTION OF RECOMBINANT PROTEINS IN VIVO AND USE FOR GENERATIN
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 13062-004001
; CURRENT APPLICATION NUMBER: US/10/077,210
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/318,474
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence
US-10-077-210-4

Alignment Scores:

Pred. No.:	178	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-077-210-4 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 64
US-10-608-710-9
; Sequence 9, Application US/10608710
; Publication No. US20040117863A1
; GENERAL INFORMATION:
; APPLICANT: GTC Biotherapeutics, Inc.
; APPLICANT: Edge, Michael D
; APPLICANT: Pollock, Daniel
; APPLICANT: Echelard, Yann
; APPLICANT: Meade, Harry M
; APPLICANT: Rybak, Susanna M
; TITLE OF INVENTION: Transgenically Produced Fusion Proteins
; FILE REFERENCE: GTC-42D
; CURRENT APPLICATION NUMBER: US/10/608,710
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 09/398,610
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 11

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: peptide linker
US-10-608-710-9

Alignment Scores:
Pred. No.:              178           Length:         17
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-608-710-9 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 65
US-10-768-873-4
; Sequence 4, Application US/10768873
; Publication No. US20040143857A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Michael
;  APPLICANT:  Meade, Harry
;  APPLICANT:  Krane, Ian
;  TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
;  FILE REFERENCE: 10275/041001
;  CURRENT APPLICATION NUMBER: US/10/768,873
;  CURRENT FILING DATE:  2004-01-30
;  PRIOR APPLICATION NUMBER: US/09/333,213
;  PRIOR FILING DATE: 1999-06-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated linker sequence
US-10-768-873-4

Alignment Scores:
Pred. No.:              178           Length:         17
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     4             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-768-873-4 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 66
US-11-016-189-4
; Sequence 4, Application US/11016189
; Publication No. US20050158832A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Michael
;  APPLICANT:  Meade, Harry
;  APPLICANT:  Krane, Ian
;  TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
;  FILE REFERENCE: 10275/041001
;  CURRENT APPLICATION NUMBER: US/11/016,189
;  CURRENT FILING DATE:  2004-12-17
;  PRIOR APPLICATION NUMBER: US 09/333,213
;  PRIOR FILING DATE: 1999-06-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated linker sequence
US-11-016-189-4

Alignment Scores:
Pred. No.:              178           Length:         17
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     6             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-016-189-4 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 67
US-11-049-853-4
; Sequence 4, Application US/11049853
; Publication No. US20050181482A1
; GENERAL INFORMATION:
;  APPLICANT: Meade, Harry
;  APPLICANT:  Krane, Ian
;  TITLE OF INVENTION: A METHOD FOR THE PRODUCTION OF AN ERYTHROPOIETIN
;  TITLE OF INVENTION:  ANALOG-HUMAN IgG FUSION PROTEINS IN TRANSGENIC MAMMAL MILK
;  FILE REFERENCE: GTC-222
;  CURRENT APPLICATION NUMBER: US/11/049,853
;  CURRENT FILING DATE:  2005-02-03
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated linker sequence
US-11-049-853-4

Alignment Scores:
Pred. No.:              178           Length:         17
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     6             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-049-853-4 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 68
US-10-283-017-2087
; Sequence 2087, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Durham, Margarita
```

```
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C20
;  CURRENT APPLICATION NUMBER: US/10/283,017
;  CURRENT FILING DATE:  2002-10-28
;  NUMBER OF SEQ ID NOS: 2157
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2087
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-283-017-2087

Alignment Scores:
Pred. No.:                174             Length:         20
Score:                    9.00            Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.7%            Indels:         0
DB:                       4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-283-017-2087 (1-20)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||
Db          8 SerSerHisSerSerSerSerSerSer 16

RESULT 69
US-10-283-017-2088
; Sequence 2088, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C20
;  CURRENT APPLICATION NUMBER: US/10/283,017
;  CURRENT FILING DATE:  2002-10-28
;  NUMBER OF SEQ ID NOS: 2157
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2088
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-283-017-2088

Alignment Scores:
Pred. No.:                174             Length:         20
Score:                    9.00            Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.7%            Indels:         0
DB:                       4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-283-017-2088 (1-20)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              |||||||||||||||||||||||||||
Db          3 SerSerHisSerSerSerSerSerSer 11

RESULT 70
US-09-996-069-12
; Sequence 12, Application US/09996069
; Publication No. US20030036199A1
; GENERAL INFORMATION:
;  APPLICANT: Bamdad, Cynthia
;  APPLICANT:  Bamdad, R. Shoshana
;  TITLE OF INVENTION: DIAGNOSTIC TUMOR MARKERS, DRUG SCREENING FOR TUMORIGENESIS INHIB
;  TITLE OF INVENTION:  AND COMPOSITIONS AND METHODS FOR TREATMENT OF CANCER
;  FILE REFERENCE: M01015/70071
;  CURRENT APPLICATION NUMBER: US/09/996,069
;  CURRENT FILING DATE:  2001-11-27
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-996-069-12

Alignment Scores:
Pred. No.:                164             Length:         31
Score:                    9.00            Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.7%            Indels:         0
DB:                       3               Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-996-069-12 (1-31)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db         10 SerGlySerSerSerSerGlySerSer 18

RESULT 71
US-10-003-681-1
; Sequence 1, Application US/10003681
; Publication No. US20020156112A1
; GENERAL INFORMATION:
;  APPLICANT: Bamdad, R. Shoshana
;  APPLICANT:  Bamdad, Cynthia
;  TITLE OF INVENTION: ENDOSTATIN-LIKE ANGIOGENISIS INHIBITION
;  FILE REFERENCE: M01015.70046.US
;  CURRENT APPLICATION NUMBER: US/10/003,681
;  CURRENT FILING DATE:  2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/248,865
;  PRIOR FILING DATE: 2000-11-15
;  PRIOR APPLICATION NUMBER: US 60/277,922
;  PRIOR FILING DATE: 2001-03-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Synthetic sequence
US-10-003-681-1

Alignment Scores:
Pred. No.:                164             Length:         31
Score:                    9.00            Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              1.7%            Indels:         0
DB:                       4               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-003-681-1 (1-31)
```

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db         10 SerGlySerSerSerSerGlySerSer 18

RESULT 72
US-10-818-694-19
; Sequence 19, Application US/10818694
; Publication No. US20050089931A1
; GENERAL INFORMATION:
;  APPLICANT: Bamdad, Cynthia C.
;  TITLE OF INVENTION: THE USE OF SELF-ASSEMBLED MONOLAYERS TO
;  TITLE OF INVENTION:  PROBE THE STRUCTURE OF A TARGET MOLECULE
;  FILE REFERENCE: M1015/7004/TJO
;  CURRENT APPLICATION NUMBER: US/10/818,694
;  CURRENT FILING DATE:  2004-04-06
;  PRIOR APPLICATION NUMBER: U.S. 60/087,766
;  PRIOR FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 1X-linker-1X
US-10-818-694-19

Alignment Scores:
Pred. No.:              164          Length:        31
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     5            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-818-694-19 (1-31)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db         12 SerGlySerSerSerSerGlySerSer 20

RESULT 73
US-10-425-115-312386
; Sequence 312386, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 312386
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(91)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_47952C.1.pep
US-10-425-115-312386

Alignment Scores:
Pred. No.:              141          Length:        91
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-425-115-312386 (1-91)

Qy        457 TCTCATTCGAGCAGCAGCAGCAGCTTT 483
              ||||||||||||||||||||||||||
Db         34 SerHisSerSerSerSerSerSerPhe 42

RESULT 74
US-10-767-701-48195
; Sequence 48195, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48195
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(99)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3481-018-P1-K1-C6.pep
US-10-767-701-48195

Alignment Scores:
Pred. No.:              139          Length:        99
Score:                  9.00         Matches:       9
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.7%         Indels:        0
DB:                     4            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-767-701-48195 (1-99)

Qy       1163 AGGGACCCTGCTCTCCCTCCAGTTCTC 1189
              ||||||||||||||||||||||||||
Db         25 ArgAspProAlaLeuProProValLeu 33

RESULT 75
US-10-425-115-302039
; Sequence 302039, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 302039
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_38534C.1.pep
US-10-425-115-302039

Alignment Scores:
```

Pred. No.:	139	Length:	99
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	4	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-425-115-302039 (1-99)

```
Qy        759 GGTGGTGGTGGTGGACCAGCACGGTTC 785
              ||||||||||||||||||||||||||
Db         64 GlyGlyGlyGlyGlyProAlaArgPhe 72
```

Search completed: March 21, 2006, 22:32:46
Job time : 172.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:28:14 ; Search time 2.8 Seconds
(without alignments)
3250.742 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 169630 seqs, 28622889 residues

Word size: 1

Total number of hits satisfying chosen parameters: 303540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10677865/runat_21032006_152113_27001/app_query.fasta_1
-DB=Published_Applications_AA_New -QFMT=fastan -SUFFIX=oligo.rapbn
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=50 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss05h
-USER=US10677865_@CGN_1_1_8_@runat_21032006_152113_27001 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	10	1.9	406	7	US-11-072-512-3291	Sequence 3291, Ap
2	9	1.7	17	7	US-11-019-955-19	Sequence 19, Appl
3	9	1.7	17	7	US-11-166-994-9	Sequence 9, Appli
4	9	1.7	118	7	US-11-156-084-63	Sequence 63, Appl
5	9	1.7	181	7	US-11-156-084-66	Sequence 66, Appl
6	9	1.7	212	6	US-10-330-773-286	Sequence 286, App
7	9	1.7	265	7	US-11-096-568A-23105	Sequence 23105, A
8	9	1.7	278	7	US-11-195-739-9	Sequence 9, Appli
9	9	1.7	332	7	US-11-096-568A-23104	Sequence 23104, A
10	9	1.7	357	7	US-11-156-084-53	Sequence 53, Appl
11	9	1.7	357	7	US-11-156-084-59	Sequence 59, Appl
12	9	1.7	357	7	US-11-156-084-169	Sequence 169, App
c 13	9	1.7	369	7	US-11-096-568A-17040	Sequence 17040, A
c 14	9	1.7	388	7	US-11-096-568A-17039	Sequence 17039, A
c 15	9	1.7	418	7	US-11-096-568A-22303	Sequence 22303, A
c 16	9	1.7	420	7	US-11-096-568A-22302	Sequence 22302, A
17	9	1.7	422	7	US-11-096-568A-33852	Sequence 33852, A
18	9	1.7	430	6	US-10-517-939-80	Sequence 80, Appl
19	9	1.7	440	6	US-10-063-703-52	Sequence 52, Appl
20	9	1.7	440	7	US-11-102-240-52	Sequence 52, Appl
21	9	1.7	440	7	US-11-103-195-52	Sequence 52, Appl
c 22	9	1.7	444	7	US-11-096-568A-22301	Sequence 22301, A
23	9	1.7	568	7	US-11-096-568A-33851	Sequence 33851, A
24	9	1.7	577	6	US-10-623-155-500	Sequence 500, App
25	9	1.7	590	7	US-11-096-568A-33850	Sequence 33850, A
26	9	1.7	802	7	US-11-087-099-4882	Sequence 4882, Ap
27	9	1.7	823	7	US-11-087-099-1971	Sequence 1971, Ap
28	9	1.7	823	7	US-11-087-099-11109	Sequence 11109, A
29	8	1.5	54	7	US-11-150-054A-17	Sequence 17, Appl
30	8	1.5	107	7	US-11-072-512-2856	Sequence 2856, Ap
c 31	8	1.5	109	6	US-10-515-241-3	Sequence 3, Appli
32	8	1.5	113	6	US-10-821-234-1406	Sequence 1406, Ap
c 33	8	1.5	153	7	US-11-096-568A-10321	Sequence 10321, A
34	8	1.5	154	6	US-10-821-234-1344	Sequence 1344, Ap
35	8	1.5	161	7	US-11-072-512-2931	Sequence 2931, Ap
c 36	8	1.5	183	7	US-11-096-568A-19290	Sequence 19290, A
37	8	1.5	216	7	US-11-096-568A-6737	Sequence 6737, Ap
c 38	8	1.5	225	7	US-11-096-568A-19289	Sequence 19289, A
39	8	1.5	251	7	US-11-096-568A-5784	Sequence 5784, Ap
40	8	1.5	253	6	US-10-724-598-28	Sequence 28, Appl
41	8	1.5	255	7	US-11-096-568A-32032	Sequence 32032, A
42	8	1.5	258	7	US-11-072-512-3034	Sequence 3034, Ap
43	8	1.5	258	7	US-11-096-568A-32031	Sequence 32031, A
44	8	1.5	274	7	US-11-087-099-10097	Sequence 10097, A
45	8	1.5	286	6	US-10-714-887-196	Sequence 196, App
46	8	1.5	286	7	US-11-096-568A-5783	Sequence 5783, Ap
47	8	1.5	303	6	US-10-517-939-214	Sequence 214, App
c 48	8	1.5	303	7	US-11-198-819-2	Sequence 2, Appli
49	8	1.5	305	7	US-11-072-512-2670	Sequence 2670, Ap
50	8	1.5	307	6	US-10-510-941-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-11-072-512-3291
; Sequence 3291, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978

```
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3291
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-3291

Alignment Scores:
Pred. No.:              0.564          Length:         406
Score:                  10.00          Matches:        10
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            1.9%           Indels:         0
DB:                     7              Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-072-512-3291 (1-406)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81

RESULT 2
US-11-019-955-19
; Sequence 19, Application US/11019955
; Publication No. US20050282763A1
; GENERAL INFORMATION:
;  APPLICANT: Hedley, Mary Lynne
;  TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
;  FILE REFERENCE: 08191-022001
;  CURRENT APPLICATION NUMBER: US/11/019,955
;  CURRENT FILING DATE:  2004-12-22
;  PRIOR APPLICATION NUMBER: US/10/074,956
;  PRIOR FILING DATE: 2002-02-12
;  PRIOR APPLICATION NUMBER: 60/268,175
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-019-955-19

Alignment Scores:
Pred. No.:              7.73           Length:         17
Score:                  9.00           Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            1.7%           Indels:         0
DB:                     7              Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-019-955-19 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 3
US-11-166-994-9
; Sequence 9, Application US/11166994
; Publication No. US20060026695A1
; GENERAL INFORMATION:
;  APPLICANT: GTC Biotherapeutics, Inc.
;  APPLICANT:  Edge, Michael D
;  APPLICANT:  Pollock, Daniel
;  APPLICANT:  Echelard, Yann
;  APPLICANT:  Meade, Harry M
;  APPLICANT:  Rybak, Susanna M
;  TITLE OF INVENTION: Transgenically Produced Fusion Proteins
;  FILE REFERENCE: GTC-42D
;  CURRENT APPLICATION NUMBER: US/11/166,994
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: US/10/608,710
;  PRIOR FILING DATE: 2003-06-27
;  PRIOR APPLICATION NUMBER: US 09/398,610
;  PRIOR FILING DATE: 1999-09-17
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: peptide linker
US-11-166-994-9

Alignment Scores:
Pred. No.:              7.73           Length:         17
Score:                  9.00           Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            1.7%           Indels:         0
DB:                     7              Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-166-994-9 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 4
US-11-156-084-63
; Sequence 63, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology LLC
;  TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
;  TITLE OF INVENTION:  agronomically interesting phenotypes
;  FILE REFERENCE: (38-21)
;  CURRENT APPLICATION NUMBER: US/11/156,084
;  CURRENT FILING DATE:  2005-06-17
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (88)..(88)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-11-156-084-63

Alignment Scores:
Pred. No.:              5.67           Length:         118
Score:                  9.00           Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            1.7%           Indels:         0
DB:                     7              Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-156-084-63 (1-118)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35

RESULT 5
US-11-156-084-66
; Sequence 66, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology LLC
;  TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
```

; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-156-084-66

Alignment Scores:

Pred. No.:	5.3	Length:	181
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-156-084-66 (1-181)

```
Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35
```

RESULT 6
US-10-330-773-286
; Sequence 286, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-773-286

Alignment Scores:

Pred. No.:	5.17	Length:	212
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-330-773-286 (1-212)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        117 SerSerGlySerSerSerGlySerSer 125
```

RESULT 7
US-11-096-568A-23105
; Sequence 23105, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23105
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(265)
; OTHER INFORMATION: Ceres Seq. ID no. 12411286
US-11-096-568A-23105

Alignment Scores:

Pred. No.:	4.99	Length:	265
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-23105 (1-265)

```
Qy        759 GGTGGTGGTGGTGGACCAGCACGGTTC 785
              ||||||||||||||||||||||||||
Db        230 GlyGlyGlyGlyGlyProAlaArgPhe 238
```

RESULT 8
US-11-195-739-9
; Sequence 9, Application US/11195739
; Publication No. US20060019324A1
; GENERAL INFORMATION:
; APPLICANT: ALZARI, PEDRO
; APPLICANT: BOITEL, BIRGITTE
; APPLICANT: VILLARINO, ANDREA
; APPLICANT: FERNANDEZ, PABLO
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: PKNB KINASE AND PSTP PHOSPHATASE AND METHODS OF IDENTIFYING
; TITLE OF INVENTION: INHIBITORY SUBSTANCES
; FILE REFERENCE: 252853US
; CURRENT APPLICATION NUMBER: US/11/195,739
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US/10/892,170
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: US 60/487,943
; PRIOR FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Bifidobacterium longum
US-11-195-739-9

Alignment Scores:

Pred. No.:	4.95	Length:	278
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-195-739-9 (1-278)

```
Qy        277 AGCGGATCCAGCATTGCCCAGGGTGGT 303
              |||||||||||||||||||||||||||
Db        187 SerGlySerSerIleAlaGlnGlyGly 195
```

RESULT 9
US-11-096-568A-23104
; Sequence 23104, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A

; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23104
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: Ceres Seq. ID no. 12411285
US-11-096-568A-23104

Alignment Scores:

Pred. No.:	4.81	Length:	332
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-23104 (1-332)

```
Qy        759 GGTGGTGGTGGTGGACCAGCACGGTTC 785
              ||||||||||||||||||||||||||
Db        297 GlyGlyGlyGlyGlyProAlaArgPhe 305
```

RESULT 10
US-11-156-084-53
; Sequence 53, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-156-084-53

Alignment Scores:

Pred. No.:	4.75	Length:	357
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-156-084-53 (1-357)

```
Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35
```

RESULT 11
US-11-156-084-59
; Sequence 59, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-156-084-59

Alignment Scores:

Pred. No.:	4.75	Length:	357
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-156-084-59 (1-357)

```
Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35
```

RESULT 12
US-11-156-084-169
; Sequence 169, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 169
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-156-084-169

Alignment Scores:

Pred. No.:	4.75	Length:	357
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-156-084-169 (1-357)

```
Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db         27 SerSerHisSerSerSerSerSerSer 35
```

RESULT 13
US-11-096-568A-17040
; Sequence 17040, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 17040
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(369)
; OTHER INFORMATION: Ceres Seq. ID no. 12355879
US-11-096-568A-17040

Alignment Scores:

Pred. No.:	4.73	Length:	369
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-17040 (1-369)

```
Qy        778 GCTGGTCCACCACCACCACCACAGGCC 752
              |||||||||||||||||||||||||||
Db         17 AlaGlyProProProProProGlnAla 25
```

RESULT 14
US-11-096-568A-17039
; Sequence 17039, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 17039
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(388)
; OTHER INFORMATION: Ceres Seq. ID no. 12355878
US-11-096-568A-17039

Alignment Scores:

Pred. No.:	4.69	Length:	388
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-17039 (1-388)

```
Qy        778 GCTGGTCCACCACCACCACCACAGGCC 752
              |||||||||||||||||||||||||||
Db         36 AlaGlyProProProProProGlnAla 44
```

RESULT 15
US-11-096-568A-22303
; Sequence 22303, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22303
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(418)
; OTHER INFORMATION: Ceres Seq. ID no. 12408439
US-11-096-568A-22303

Alignment Scores:

Pred. No.:	4.64	Length:	418
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-22303 (1-418)

```
Qy        778 GCTGGTCCACCACCACCACCACAGGCC 752
              |||||||||||||||||||||||||||
Db         66 AlaGlyProProProProProGlnAla 74
```

RESULT 16
US-11-096-568A-22302
; Sequence 22302, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22302
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(420)
; OTHER INFORMATION: Ceres Seq. ID no. 12408438
US-11-096-568A-22302

Alignment Scores:

Pred. No.:	4.63	Length:	420
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-22302 (1-420)

```
Qy        778 GCTGGTCCACCACCACCACCACAGGCC 752
              |||||||||||||||||||||||||||
Db         68 AlaGlyProProProProProGlnAla 76
```

RESULT 17
US-11-096-568A-33852
; Sequence 33852, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33852
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(422)
; OTHER INFORMATION: Ceres Seq. ID no. 13604334
US-11-096-568A-33852

Alignment Scores:

```
Pred. No.:              4.63          Length:        422
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-33852 (1-422)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        374 SerGlySerSerSerSerGlySerSer 382

RESULT 18
US-10-517-939-80
; Sequence 80, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
;  APPLICANT: Steer, Brian
;  APPLICANT:  Callen, Walter
;  APPLICANT:  Healey, Shaun
;  APPLICANT:  Hazlewood, Geoff
;  APPLICANT:  Wu, Di
;  APPLICANT:  Blum, David
;  APPLICANT:  Esteghlalian, Alireza
;  TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
;  TITLE OF INVENTION:  AND METHODS FOR MAKING AND USING THEM
;  FILE REFERENCE: 564462007901
;  CURRENT APPLICATION NUMBER: US/10/517,939
;  CURRENT FILING DATE:  2004-12-13
;  PRIOR APPLICATION NUMBER: PCT/US03/19153
;  PRIOR FILING DATE: 2003-06-16
;  PRIOR APPLICATION NUMBER: 60/389,299
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 380
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 430
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Obtained from an environmental sample
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: (1)...(36)
US-10-517-939-80

Alignment Scores:
Pred. No.:              4.62          Length:        430
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-517-939-80 (1-430)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||||
Db        341 SerSerSerSerSerSerGlnSerSer 349

RESULT 19
US-10-063-703-52
; Sequence 52, Application US/10063703
; Publication No. US20060008901A1
; GENERAL INFORMATION:
;  APPLICANT: Eaton,Dan L.
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3230R1C1
;  CURRENT APPLICATION NUMBER: US/10/063,703
;  CURRENT FILING DATE:  2002-05-08
;      Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-063-703-52

Alignment Scores:
Pred. No.:              4.6           Length:        440
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-063-703-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 20
US-11-102-240-52
; Sequence 52, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
;  APPLICANT: Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRES
;  TITLE OF INVENTION:  ESOPHAGEAL TUMOR
;  FILE REFERENCE: P3230R1C106C
;  CURRENT APPLICATION NUMBER: US/11/102,240
;  CURRENT FILING DATE:  2005-04-08
;  PRIOR APPLICATION NUMBER: 10/063662
;  PRIOR FILING DATE: 2002-05-07
;  PRIOR APPLICATION NUMBER: 10/006867
;  PRIOR FILING DATE: 2001-12-06
;  PRIOR APPLICATION NUMBER: PCT/US00/23328
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: 60/170262
;  PRIOR FILING DATE: 199-12-09
;  NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-11-102-240-52

Alignment Scores:
Pred. No.:              4.6           Length:        440
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-102-240-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281
```

RESULT 21
US-11-103-195-52
; Sequence 52, Application US/11103195
; Publication No. US20060051779A1
; GENERAL INFORMATION:
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230C166C1
; CURRENT APPLICATION NUMBER: US/11/103,195
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: 10/063743
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-103-195-52

Alignment Scores:

Pred. No.:	4.6	Length:	440
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-103-195-52 (1-440)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281
```

RESULT 22
US-11-096-568A-22301
; Sequence 22301, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22301
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(444)
; OTHER INFORMATION: Ceres Seq. ID no. 12408437
US-11-096-568A-22301

Alignment Scores:

Pred. No.:	4.59	Length:	444
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-22301 (1-444)

```
Qy        778 GCTGGTCCACCACCACCACCACAGGCC 752
              ||||||||||||||||||||||||||||
Db         92 AlaGlyProProProProProGlnAla 100
```

RESULT 23
US-11-096-568A-33851
; Sequence 33851, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33851
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(568)
; OTHER INFORMATION: Ceres Seq. ID no. 13604333
US-11-096-568A-33851

Alignment Scores:

Pred. No.:	4.41	Length:	568
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-33851 (1-568)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        520 SerGlySerSerSerSerGlySerSer 528
```

RESULT 24
US-10-623-155-500
; Sequence 500, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C20
; CURRENT APPLICATION NUMBER: US/10/623,155
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-155-500

Alignment Scores:

Pred. No.:	4.4	Length:	577
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0

```
DB:                    6              Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-623-155-500 (1-577)

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAG 1521
              ||||||||||||||||||||||||||||
Db        553 IleArgAspIleLeuAlaGlnValLys 561

RESULT 25
US-11-096-568A-33850
; Sequence 33850, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33850
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(590)
;   OTHER INFORMATION: Ceres Seq. ID no. 13604332
US-11-096-568A-33850

Alignment Scores:
Pred. No.:              4.39          Length:        590
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-33850 (1-590)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        542 SerGlySerSerSerSerGlySerSer 550

RESULT 26
US-11-087-099-4882
; Sequence 4882, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4882
;   LENGTH: 802
;   TYPE: PRT
;   ORGANISM: Phanerochaete chrysosporium
US-11-087-099-4882

Alignment Scores:
Pred. No.:              4.18          Length:        802
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-087-099-4882 (1-802)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||
Db         45 SerSerSerHisSerSerSerSerSer 53

RESULT 27
US-11-087-099-1971
; Sequence 1971, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1971
;   LENGTH: 823
;   TYPE: PRT
;   ORGANISM: Phanerochaete chrysosporium
US-11-087-099-1971

Alignment Scores:
Pred. No.:              4.16          Length:        823
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-087-099-1971 (1-823)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||
Db         66 SerSerSerHisSerSerSerSerSer 74

RESULT 28
US-11-087-099-11109
; Sequence 11109, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11109
;   LENGTH: 823
;   TYPE: PRT
;   ORGANISM: Phanerochaete chrysosporium
US-11-087-099-11109

Alignment Scores:
Pred. No.:              4.16          Length:        823
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-087-099-11109 (1-823)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGCAGC 477
              ||||||||||||||||||||||||||||
Db         66 SerSerSerHisSerSerSerSerSer 74

RESULT 29
US-11-150-054A-17
; Sequence 17, Application US/11150054A
; Publication No. US20050278801A1
; GENERAL INFORMATION:
;  APPLICANT: Lassner, Michael
;  APPLICANT:  Wilkinson, Jack Q.
;  TITLE OF INVENTION: Plastid Transit Peptides
;  FILE REFERENCE: 2119-4284US1
```

```
;  CURRENT APPLICATION NUMBER: US/11/150,054A
;  CURRENT FILING DATE:  2005-06-09
;  PRIOR APPLICATION NUMBER: 60/578,535
;  PRIOR FILING DATE: 2004-06-09
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 17
;   LENGTH: 54
;   TYPE: PRT
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
US-11-150-054A-17

Alignment Scores:
Pred. No.:              53          Length:         54
Score:                  8.00        Matches:        8
Percent Similarity:     100.0%      Conservative:   0
Best Local Similarity:  100.0%      Mismatches:     0
Query Match:            1.5%        Indels:         0
DB:                     7           Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-150-054A-17 (1-54)

Qy        583 AGCTCTTCCTCTTCCCAAACCTCT 606
              ||||||||||||||||||||||||
Db         22 SerSerSerSerSerGlnThrSer 29

RESULT 30
US-11-072-512-2856
; Sequence 2856, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2856
;   LENGTH: 107
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2856

Alignment Scores:
Pred. No.:              47.5        Length:         107
Score:                  8.00        Matches:        8
Percent Similarity:     100.0%      Conservative:   0
Best Local Similarity:  100.0%      Mismatches:     0
Query Match:            1.5%        Indels:         0
DB:                     7           Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-072-512-2856 (1-107)

Qy       1297 AGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||
Db         10 SerSerSerSerSerGlnSerSer 17

RESULT 31
US-10-515-241-3
; Sequence 3, Application US/10515241
; Publication No. US20060035320A1
; GENERAL INFORMATION:
;  APPLICANT: ESBATECH AG
;  TITLE OF INVENTION: Immunoglobulin Frameworks which demonstrate enhanced stability i
;  TITLE OF INVENTION:  intracellular environment and methods of identifying same
;  FILE REFERENCE: 08420PC
;  CURRENT APPLICATION NUMBER: US/10/515,241
;  CURRENT FILING DATE:  2004-11-19
;  PRIOR APPLICATION NUMBER: US 60/382,649
;  PRIOR FILING DATE: 2002-05-22
;  PRIOR APPLICATION NUMBER: US 60/438,246
;  PRIOR FILING DATE: 2003-01-03
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 109
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: antibody framework
US-10-515-241-3

Alignment Scores:
Pred. No.:              47.4        Length:         109
Score:                  8.00        Matches:        8
Percent Similarity:     100.0%      Conservative:   0
Best Local Similarity:  100.0%      Mismatches:     0
Query Match:            1.5%        Indels:         0
DB:                     6           Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-515-241-3 (1-109)

Qy        264 GGAGCCACCACCAGAGCTTCTGGC 241
              ||||||||||||||||||||||||
Db         50 GlyAlaThrThrArgAlaSerGly 57

RESULT 32
US-10-821-234-1406
; Sequence 1406, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1406
;   LENGTH: 113
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1406

Alignment Scores:
Pred. No.:              47.1        Length:         113
Score:                  8.00        Matches:        8
```

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 6 Gaps: 0

US-10-677-865A-2 (1-1590) x US-10-821-234-1406 (1-113)

Qy 415 TCTCACTCGGGAAGCAGCGGCTCT 438
||||||||||||||||||||||||
Db 15 SerHisSerGlySerSerGlySer 22

RESULT 33
US-11-096-568A-10321
; Sequence 10321, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 10321
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(153)
; OTHER INFORMATION: Ceres Seq. ID no. 13590886
US-11-096-568A-10321

Alignment Scores:
Pred. No.: 44.9 Length: 153
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 7 Gaps: 0

US-10-677-865A-2 (1-1590) x US-11-096-568A-10321 (1-153)

Qy 1302 GCTGCTGCTGAAGGAGCCGGTGCC 1279
||||||||||||||||||||||||
Db 49 AlaAlaAlaGluGlyAlaGlyAla 56

RESULT 34
US-10-821-234-1344
; Sequence 1344, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1344
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1344

Alignment Scores:
Pred. No.: 44.9 Length: 154
Score: 8.00 Matches: 8

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 6 Gaps: 0

US-10-677-865A-2 (1-1590) x US-10-821-234-1344 (1-154)

Qy 1445 GTGCCAAGCCCTGTGGCTCCAGCA 1468
||||||||||||||||||||||||
Db 129 ValProSerProValAlaProAla 136

RESULT 35
US-11-072-512-2931
; Sequence 2931, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2931
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2931

Alignment Scores:
Pred. No.: 44.5 Length: 161
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.5% Indels: 0
DB: 7 Gaps: 0

US-10-677-865A-2 (1-1590) x US-11-072-512-2931 (1-161)

Qy 55 GCACTGCTGCTGGCTGGTCTCCTC 78
||||||||||||||||||||||||
Db 149 AlaLeuLeuLeuAlaGlyLeuLeu 156

RESULT 36
US-11-096-568A-19290
; Sequence 19290, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2

```
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19290
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(183)
;   OTHER INFORMATION: Ceres Seq. ID no. 12372543
US-11-096-568A-19290

Alignment Scores:
Pred. No.:              43.6          Length:         183
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-19290 (1-183)

Qy         86 CCTGGCAGGAGGAGACCAGCCAGC 63
              ||||||||||||||||||||||||
Db         15 ProGlyArgArgArgProAlaSer 22

RESULT 37
US-11-096-568A-6737
; Sequence 6737, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6737
;   LENGTH: 216
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(216)
;   OTHER INFORMATION: Ceres Seq. ID no. 14316456
US-11-096-568A-6737

Alignment Scores:
Pred. No.:              42.5          Length:         216
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-6737 (1-216)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         26 SerPheSerSerSerSerSerSer 33

RESULT 38
US-11-096-568A-19289
; Sequence 19289, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19289
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(225)
;   OTHER INFORMATION: Ceres Seq. ID no. 12372542
US-11-096-568A-19289

Alignment Scores:
Pred. No.:              42.2          Length:         225
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-19289 (1-225)

Qy         86 CCTGGCAGGAGGAGACCAGCCAGC 63
              ||||||||||||||||||||||||
Db         57 ProGlyArgArgArgProAlaSer 64

RESULT 39
US-11-096-568A-5784
; Sequence 5784, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5784
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(251)
;   OTHER INFORMATION: Ceres Seq. ID no. 14311779
US-11-096-568A-5784

Alignment Scores:
Pred. No.:              41.5          Length:         251
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-5784 (1-251)

Qy         400 CTGGGGAGCAGCAGCTCTCACTCG 423
               ||||||||||||||||||||||||
Db          71 LeuGlySerSerSerSerHisSer 78

RESULT 40
US-10-724-598-28
; Sequence 28, Application US/10724598
; Publication No. US20060019352A1
;   GENERAL INFORMATION:
;       APPLICANT: BLANCHE, FRANCIS; CAMERON, BEATRICE; CROUZET,
;                  JOEL;DEBUSSCHE, LAURENT; LEVCY SCHIL, SOPHIE;
;                  THIBAUT, DENIS
;       TITLE OF INVENTION: POLYPEPTIDES INVOLVED IN THE
;                           BIOSYNTHESIS OF COBALAMINS AND/OR COBAMIDES, DNA SEQUENCES
;                           CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEI
```

```
;                          USE.
;       NUMBER OF SEQUENCES: 60
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: MORGAN & FINNEGAN
;            STREET: 555 13TH STREET, N.W.
;            CITY: WASHINGTON
;            STATE: DISTRICT OF COLUMBIA
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: FLOPPY DISK
;            COMPUTER: IBM PC COMPATIBLE
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: WORDPERFECT 5.1
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/724,598
;            FILING DATE: 01-Dec-2003
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/916,151
;            FILING DATE: 14-SEP-1992
;            APPLICATION NUMBER: PCT/FR91/00054
;            FILING DATE: 30-JAN-1991
;       ATTORNEY/AGENT INFORMATION:
;            NAME: F. F. CALVETTI
;            REGISTRATION NUMBER: 28,557
;            REFERENCE/DOCKET NUMBER: 1290-7213
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 857-7887
;            TELEFAX: (202) 857-7929
;   INFORMATION FOR SEQ ID NO: 28:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 253 amino acids
;            TYPE: Amino Acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: No
;       ORIGINAL SOURCE:
;            ORGANISM: Pseudomonas denitrificans
;            STRAIN: <Unknown>
;            INDIVIDUAL ISOLATE: <Unknown>
;            DEVELOPMENTAL STAGE: <Unknown>
;            HAPLOTYPE: <Unknown>
;            TISSUE TYPE: <Unknown>
;            CELL TYPE: <Unknown>
;            CELL LINE: <Unknown>
;            ORGANELLE: <Unknown>
;       FEATURE:
;            NAME/KEY:  COBM
;            LOCATION:
;            IDENTIFICATION METHOD:
;            OTHER INFORMATION:  Translation product of SEQ ID NO:27
;       SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-10-724-598-28

Alignment Scores:
Pred. No.:              41.5          Length:        253
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-724-598-28 (1-253)

Qy       1000 CCAGGGGTCCCTTCCTTTGCAGCT 1023
              ||||||||||||||||||||||||
Db        108 ProGlyValProSerPheAlaAla 115

RESULT 41
US-11-096-568A-32032
; Sequence 32032, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32032
;   LENGTH: 255
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(255)
;   OTHER INFORMATION: Ceres Seq. ID no. 13591997
US-11-096-568A-32032

Alignment Scores:
Pred. No.:              41.4          Length:        255
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-32032 (1-255)

Qy        457 TCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||
Db         97 SerHisSerSerSerSerSerSer 104

RESULT 42
US-11-072-512-3034
; Sequence 3034, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3034
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-3034

Alignment Scores:
```

```
Pred. No.:              41.3          Length:        258
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-072-512-3034 (1-258)

Qy        451 AGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||
Db         66 SerSerSerHisSerSerSerSer 73

RESULT 43
US-11-096-568A-32031
; Sequence 32031, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32031
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(258)
;   OTHER INFORMATION: Ceres Seq. ID no. 13591996
US-11-096-568A-32031

Alignment Scores:
Pred. No.:              41.3          Length:        258
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-32031 (1-258)

Qy        457 TCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||
Db        100 SerHisSerSerSerSerSerSer 107

RESULT 44
US-11-087-099-10097
; Sequence 10097, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10097
;   LENGTH: 274
;   TYPE: PRT
;   ORGANISM: Lycopersicon esculentum
US-11-087-099-10097

Alignment Scores:
Pred. No.:              40.9          Length:        274
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-087-099-10097 (1-274)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db        235 SerPheSerSerSerSerSerSer 242

RESULT 45
US-10-714-887-196
; Sequence 196, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
;  APPLICANT: Mendel Biotechnology, Inc.
;  APPLICANT:  HEARD, Jacqueline
;  APPLICANT:  RIECHMANN, Jose Luis
;  APPLICANT:  CREELMAN, Robert
;  APPLICANT:  RATCLIFFE, Oliver
;  APPLICANT:  CANALES, Roger
;  APPLICANT:  REPETTI, Peter
;  APPLICANT:  KUMIMOTO, Roderick W
;  APPLICANT:  GUTTERSON, Neal
;  APPLICANT:  REUBER, T. Lynne
;  APPLICANT:  PINEDA, Omaira
;  APPLICANT:  SHERMAN, Bradley K
;  TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
;  FILE REFERENCE: MBI0058-CIP
;  CURRENT APPLICATION NUMBER: US/10/714,887
;  CURRENT FILING DATE:  2003-11-13
;  PRIOR APPLICATION NUMBER: 10/412,699
;  PRIOR FILING DATE: 2003-04-10
;  PRIOR APPLICATION NUMBER: 09/506,720
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/135,134
;  PRIOR FILING DATE: 1999-05-20
;  PRIOR APPLICATION NUMBER: 09/394,519
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: 09/533,392
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 09/533,029
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 09/532,591
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 09/533,030
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/125,814
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: 09/713,994
;  PRIOR FILING DATE: 2000-11-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 430
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 196
;   LENGTH: 286
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: G3772 polypeptide  Orthologous to G3086
US-10-714-887-196

Alignment Scores:
Pred. No.:              40.6          Length:        286
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-714-887-196 (1-286)

Qy       1294 AGCAGCAGCTCCAGTTCCCAATCG 1317
              ||||||||||||||||||||||||
Db          8 SerSerSerSerSerSerGlnSer 15
```

```
RESULT 46
US-11-096-568A-5783
; Sequence 5783, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5783
;   LENGTH: 286
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(286)
;   OTHER INFORMATION: Ceres Seq. ID no. 14311778
US-11-096-568A-5783

Alignment Scores:
Pred. No.:              40.6          Length:         286
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-5783 (1-286)

Qy        400 CTGGGGAGCAGCAGCTCTCACTCG 423
              ||||||||||||||||||||||||
Db        106 LeuGlySerSerSerSerHisSer 113

RESULT 47
US-10-517-939-214
; Sequence 214, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
;  APPLICANT: Steer, Brian
;  APPLICANT:  Callen, Walter
;  APPLICANT:  Healey, Shaun
;  APPLICANT:  Hazlewood, Geoff
;  APPLICANT:  Wu, Di
;  APPLICANT:  Blum, David
;  APPLICANT:  Esteghlalian, Alireza
;  TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
;  TITLE OF INVENTION:  AND METHODS FOR MAKING AND USING THEM
;  FILE REFERENCE: 564462007901
;  CURRENT APPLICATION NUMBER: US/10/517,939
;  CURRENT FILING DATE:  2004-12-13
;  PRIOR APPLICATION NUMBER: PCT/US03/19153
;  PRIOR FILING DATE: 2003-06-16
;  PRIOR APPLICATION NUMBER: 60/389,299
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 380
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 214
;   LENGTH: 303
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Obtained from an environmental sample
US-10-517-939-214

Alignment Scores:
Pred. No.:              40.3          Length:         303
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     6             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-517-939-214 (1-303)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||
Db        216 SerSerGlySerSerSerGlySer 223

RESULT 48
US-11-198-819-2
; Sequence 2, Application US/11198819
; Publication No. US20050287582A1
;  GENERAL INFORMATION:
;    APPLICANT:  Adema, Gosse Jan
;    TITLE OF INVENTION:  Isolated Mammalian Monocyte Cell Genes;
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/11/198,819
;      FILING DATE:  04-AUG-2005
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,950
;      FILING DATE:  05-DEC-1997
;      APPLICATION NUMBER:  US 60/041,279
;      FILING DATE:  21-MARCH-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,181
;      FILING DATE:  16-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/032,252
;      FILING DATE:  06-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0670K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650)852-9196
;      TELEFAX:  (650)496-1204
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-11-198-819-2

Alignment Scores:
Pred. No.:              40.3          Length:         303
Score:                  8.00          Matches:        8
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.5%          Indels:         0
DB:                     7             Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-198-819-2 (1-303)

Qy        773 TCCACCACCACCACCACAGGCCTC 750
              ||||||||||||||||||||||||
Db        171 SerThrThrThrThrThrGlyLeu 178

RESULT 49
```

US-11-072-512-2670
; Sequence 2670, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2670
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2670

Alignment Scores:

Pred. No.:	40.2	Length:	305
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-072-512-2670 (1-305)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db        209 SerPheSerSerSerSerSerSer 216
```

RESULT 50
US-10-510-941-2
; Sequence 2, Application US/10510941
; Publication No. US20060040346A1
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Rasmussen, Michael Dolberg
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Olesen, Peter Bjarke
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10297.204-US
; CURRENT APPLICATION NUMBER: US/10/510,941
; CURRENT FILING DATE: 2004-10-08
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-941-2

Alignment Scores:

Pred. No.:	40.2	Length:	307
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-510-941-2 (1-307)

```
Qy        413 GCTCTCACTCGGGAAGCAGCGGCT 436
              ||||||||||||||||||||||||
Db        237 AlaLeuThrArgGluAlaAlaAla 244
```

RESULT 51
US-11-087-099-11105
; Sequence 11105, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11105
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(310)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-11105

Alignment Scores:

Pred. No.:	40.1	Length:	310
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-087-099-11105 (1-310)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCA 490
              ||||||||||||||||||||||||
Db        254 AlaAlaAlaAlaAlaPheSerSer 261
```

RESULT 52
US-11-096-568A-7337
; Sequence 7337, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7337
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 15171356
US-11-096-568A-7337

Alignment Scores:

```
Pred. No.:              40          Length:        317
Score:                  8.00        Matches:       8
Percent Similarity:     100.0%      Conservative:  0
Best Local Similarity:  100.0%      Mismatches:    0
Query Match:            1.5%        Indels:        0
DB:                     7           Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-7337 (1-317)

Qy       1170 CTGCTCTCCCTCCAGTTCTCGAGT 1193
              ||||||||||||||||||||||||
Db         17 LeuLeuSerLeuGlnPheSerSer 24

RESULT 53
US-11-096-568A-32030
; Sequence 32030, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32030
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(325)
;   OTHER INFORMATION: Ceres Seq. ID no. 13591995
US-11-096-568A-32030

Alignment Scores:
Pred. No.:              39.8        Length:        325
Score:                  8.00        Matches:       8
Percent Similarity:     100.0%      Conservative:  0
Best Local Similarity:  100.0%      Mismatches:    0
Query Match:            1.5%        Indels:        0
DB:                     7           Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-32030 (1-325)

Qy        457 TCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||
Db        167 SerHisSerSerSerSerSerSer 174

RESULT 54
US-11-096-568A-3897
; Sequence 3897, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3897
;   LENGTH: 328
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(328)
;   OTHER INFORMATION: Ceres Seq. ID no. 13594131
US-11-096-568A-3897

Alignment Scores:
Pred. No.:              39.8        Length:        328
Score:                  8.00        Matches:       8
Percent Similarity:     100.0%      Conservative:  0
Best Local Similarity:  100.0%      Mismatches:    0
Query Match:            1.5%        Indels:        0
DB:                     7           Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-3897 (1-328)

Qy       1158 GCCCGTGGAGCCGCCTCCACAGAG 1135
              ||||||||||||||||||||||||
Db        169 AlaArgGlyAlaAlaSerThrGlu 176

RESULT 55
US-11-096-568A-3896
; Sequence 3896, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3896
;   LENGTH: 334
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(334)
;   OTHER INFORMATION: Ceres Seq. ID no. 13594130
US-11-096-568A-3896

Alignment Scores:
Pred. No.:              39.7        Length:        334
Score:                  8.00        Matches:       8
Percent Similarity:     100.0%      Conservative:  0
Best Local Similarity:  100.0%      Mismatches:    0
Query Match:            1.5%        Indels:        0
DB:                     7           Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-3896 (1-334)

Qy       1158 GCCCGTGGAGCCGCCTCCACAGAG 1135
              ||||||||||||||||||||||||
Db        175 AlaArgGlyAlaAlaSerThrGlu 182

RESULT 56
US-10-821-234-1004
; Sequence 1004, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1004
;   LENGTH: 337
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1004

Alignment Scores:
Pred. No.:              39.6        Length:        337
```

```
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-821-234-1004 (1-337)

Qy        406 AGCAGCAGCTCTCACTCGGGAAGC 429
              ||||||||||||||||||||||||
Db        308 SerSerSerSerHisSerGlySer 315

RESULT 57
US-10-501-035-264
; Sequence 264, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
;  TITLE OF INVENTION:  ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
;  TITLE OF INVENTION:  AND/OR PROTEIN TYROSINE KINASE PATHWAYS
;  FILE REFERENCE: D0185 PCT
;  CURRENT APPLICATION NUMBER: US/10/501,035
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: US 60/350,061
;  PRIOR FILING DATE: 2002-01-18
;  NUMBER OF SEQ ID NOS: 795
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 264
;   LENGTH: 338
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-501-035-264

Alignment Scores:
Pred. No.:              39.6          Length:        338
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-501-035-264 (1-338)

Qy        406 AGCAGCAGCTCTCACTCGGGAAGC 429
              ||||||||||||||||||||||||
Db        309 SerSerSerSerHisSerGlySer 316

RESULT 58
US-11-096-568A-25786
; Sequence 25786, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 25786
;   LENGTH: 341
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(341)
;   OTHER INFORMATION: Ceres Seq. ID no. 13493530
US-11-096-568A-25786

Alignment Scores:
Pred. No.:              39.5          Length:        341
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-25786 (1-341)

Qy       1052 TCTCCAGCAACCCCATCATCCCCA 1075
              ||||||||||||||||||||||||
Db         31 SerProAlaThrProSerSerPro 38

RESULT 59
US-11-096-568A-7182
; Sequence 7182, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7182
;   LENGTH: 367
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(367)
;   OTHER INFORMATION: Ceres Seq. ID no. 15170744
US-11-096-568A-7182

Alignment Scores:
Pred. No.:              39.1          Length:        367
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-7182 (1-367)

Qy       1297 AGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||
Db        177 SerSerSerSerSerGlnSerSer 184

RESULT 60
US-11-096-568A-7181
; Sequence 7181, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7181
;   LENGTH: 381
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(381)
;   OTHER INFORMATION: Ceres Seq. ID no. 15170743
US-11-096-568A-7181

Alignment Scores:
Pred. No.:              38.8          Length:        381
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.5%            Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-7181 (1-381)

Qy       1297 AGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||
Db        191 SerSerSerSerSerGlnSerSer 198

RESULT 61
US-11-096-568A-3895
; Sequence 3895, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3895
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(388)
;   OTHER INFORMATION: Ceres Seq. ID no. 13594129
US-11-096-568A-3895

Alignment Scores:
Pred. No.:              38.7            Length:         388
Score:                  8.00            Matches:        8
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.5%            Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-096-568A-3895 (1-388)

Qy       1158 GCCCGTGGAGCCGCCTCCACAGAG 1135
              ||||||||||||||||||||||||
Db        229 AlaArgGlyAlaAlaSerThrGlu 236

RESULT 62
US-10-493-909-68
; Sequence 68, Application US/10493909
; Publication No. US20060015969A1
; GENERAL INFORMATION:
;  APPLICANT: LARRICK, JAMES W.
;  APPLICANT:  WYCOFF, KEITH L.
;  TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING TOXICITY
;  TITLE OF INVENTION:  AND PATHOGEN-MEDIATED DISEASES
;  FILE REFERENCE: 41514-20004.01
;  CURRENT APPLICATION NUMBER: US/10/493,909
;  CURRENT FILING DATE:  2004-04-26
;  PRIOR APPLICATION NUMBER: PCT/US01/13932
;  PRIOR FILING DATE: 2001-04-28
;  PRIOR APPLICATION NUMBER: 60/200,298
;  PRIOR FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-493-909-68

Alignment Scores:
Pred. No.:              38.4            Length:         406
Score:                  8.00            Matches:        8
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.5%            Indels:         0
DB:                     6               Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-493-909-68 (1-406)

Qy         55 GCACTGCTGCTGGCTGGTCTCCTC 78
              ||||||||||||||||||||||||
Db          6 AlaLeuLeuLeuAlaGlyLeuLeu 13

RESULT 63
US-11-107-028-7
; Sequence 7, Application US/11107028
; Publication No. US20050276803A1
; GENERAL INFORMATION:
;  APPLICANT: CHAN, ANDREW C.
;  APPLICANT:  GONG, QIAN
;  APPLICANT:  MARTIN, FLAVIUS
;  TITLE OF INVENTION: Method for Augmenting B Cell Depletion
;  FILE REFERENCE: P2112R1
;  CURRENT APPLICATION NUMBER: US/11/107,028
;  CURRENT FILING DATE:  2005-04-15
;  PRIOR APPLICATION NUMBER: US 60/563,263
;  PRIOR FILING DATE: 2004-04-16
;  NUMBER OF SEQ ID NOS: 52
; SEQ ID NO 7
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-107-028-7

Alignment Scores:
Pred. No.:              38.4            Length:         406
Score:                  8.00            Matches:        8
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            1.5%            Indels:         0
DB:                     7               Gaps:           0

US-10-677-865A-2 (1-1590) x US-11-107-028-7 (1-406)

Qy         55 GCACTGCTGCTGGCTGGTCTCCTC 78
              ||||||||||||||||||||||||
Db          6 AlaLeuLeuLeuAlaGlyLeuLeu 13

RESULT 64
US-11-096-568A-7180
; Sequence 7180, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7180
;   LENGTH: 406
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(406)
;   OTHER INFORMATION: Ceres Seq. ID no. 15170742
US-11-096-568A-7180

Alignment Scores:
Pred. No.:              38.4            Length:         406
Score:                  8.00            Matches:        8
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
```

```
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-7180 (1-406)

Qy       1297 AGCAGCTCCAGTTCCCAATCGAGT 1320
              ||||||||||||||||||||||||
Db        216 SerSerSerSerSerGlnSerSer 223

RESULT 65
US-11-096-568A-26292
; Sequence 26292, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26292
;   LENGTH: 434
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(434)
;   OTHER INFORMATION: Ceres Seq. ID no. 13499247
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (130)..(130)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-26292

Alignment Scores:
Pred. No.:              38            Length:        434
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-26292 (1-434)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         25 SerPheSerSerSerSerSerSer 32

RESULT 66
US-11-096-568A-26291
; Sequence 26291, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26291
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(435)
;   OTHER INFORMATION: Ceres Seq. ID no. 13499246
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131)..(131)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-26291

Alignment Scores:
Pred. No.:              38            Length:        435
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     7             Gaps:          0

US-10-677-865A-2 (1-1590) x US-11-096-568A-26291 (1-435)

Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         26 SerPheSerSerSerSerSerSer 33

RESULT 67
US-10-330-773-283
; Sequence 283, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283
;   LENGTH: 460
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-330-773-283

Alignment Scores:
Pred. No.:              37.7          Length:        460
Score:                  8.00          Matches:       8
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.5%          Indels:        0
DB:                     6             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-330-773-283 (1-460)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||
Db        366 SerSerGlySerSerSerGlySer 373

RESULT 68
US-11-096-568A-26290
; Sequence 26290, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26290
;   LENGTH: 460
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(460)
;   OTHER INFORMATION: Ceres Seq. ID no. 13499245
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (156)..(156)
;   OTHER INFORMATION: Xaa is any aa, unknown or other
```

US-11-096-568A-26290

Alignment Scores:

Pred. No.:	37.7	Length:	460
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-096-568A-26290 (1-460)

```
Qy       1288 TCCTTCAGCAGCAGCTCCAGTTCC 1311
              ||||||||||||||||||||||||
Db         51 SerPheSerSerSerSerSerSer 58
```

RESULT 69
US-10-641-678-39
; Sequence 39, Application US/10641678
; Publication No. US20050277172A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony, G.
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Neefe, Paulien
; APPLICANT: Sandgren, Mats
; APPLICANT: Shaw, Andrew
; APPLICANT: Stahlberg, Jerry
; TITLE OF INVENTION: Novel Variant Hypocrea jecorina CBH1
; TITLE OF INVENTION: Cellulases
; FILE REFERENCE: GC772-3
; CURRENT APPLICATION NUMBER: US/10/641,678
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US 60/458,853
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/458,696
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/456,368
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US 60/404,063
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-10-641-678-39

Alignment Scores:

Pred. No.:	37	Length:	516
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-641-678-39 (1-516)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||
Db        464 SerSerSerHisSerSerSerSer 471
```

RESULT 70
US-10-641-678-40
; Sequence 40, Application US/10641678
; Publication No. US20050277172A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony, G.
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Neefe, Paulien
; APPLICANT: Sandgren, Mats
; APPLICANT: Shaw, Andrew
; APPLICANT: Stahlberg, Jerry
; TITLE OF INVENTION: Novel Variant Hypocrea jecorina CBH1
; TITLE OF INVENTION: Cellulases
; FILE REFERENCE: GC772-3
; CURRENT APPLICATION NUMBER: US/10/641,678
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US 60/458,853
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/458,696
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/456,368
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US 60/404,063
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-10-641-678-40

Alignment Scores:

Pred. No.:	37	Length:	516
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-641-678-40 (1-516)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||
Db        464 SerSerSerHisSerSerSerSer 471
```

RESULT 71
US-10-641-678-41
; Sequence 41, Application US/10641678
; Publication No. US20050277172A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony, G.
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Neefe, Paulien
; APPLICANT: Sandgren, Mats
; APPLICANT: Shaw, Andrew
; APPLICANT: Stahlberg, Jerry
; TITLE OF INVENTION: Novel Variant Hypocrea jecorina CBH1
; TITLE OF INVENTION: Cellulases
; FILE REFERENCE: GC772-3
; CURRENT APPLICATION NUMBER: US/10/641,678
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US 60/458,853
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/458,696
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/456,368
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US 60/404,063
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-10-641-678-41

Alignment Scores:

Pred. No.:	37	Length:	516
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-641-678-41 (1-516)

```
Qy        451 AGCAGCTCTCATTCGAGCAGCAGC 474
              ||||||||||||||||||||||||
Db        464 SerSerSerHisSerSerSerSer 471
```

RESULT 72
US-10-641-678-61
; Sequence 61, Application US/10641678
; Publication No. US20050277172A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony, G.
; APPLICANT: Goedegebuur, Frits
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Neefe, Paulien
; APPLICANT: Sandgren, Mats
; APPLICANT: Shaw, Andrew
; APPLICANT: Stahlberg, Jerry
; TITLE OF INVENTION: Novel Variant Hypocrea jecorina CBH1
; TITLE OF INVENTION: Cellulases
; FILE REFERENCE: GC772-3
; CURRENT APPLICATION NUMBER: US/10/641,678
; CURRENT FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US 60/458,853
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/458,696
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/456,368
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US 60/404,063
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-10-641-678-61

Alignment Scores:

Pred. No.:	36.8	Length:	536
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	6	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-641-678-61 (1-536)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||
Db        461 SerSerGlySerSerSerGlySer 468
```

RESULT 73
US-11-114-906-8
; Sequence 8, Application US/11114906
; Publication No. US20050266452A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0281-USA
; CURRENT APPLICATION NUMBER: US/11/114,906
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: US/10/413,437
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: US/10/020,079
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,103
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/289,422
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 751
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(751)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-114-906-8

Alignment Scores:

Pred. No.:	34.8	Length:	751
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-114-906-8 (1-751)

```
Qy        253 GGTGGTGGCTCCAGTGGTAGCTCC 276
              ||||||||||||||||||||||||
Db        689 GlyGlyGlySerSerGlySerSer 696
```

RESULT 74
US-11-114-906-6
; Sequence 6, Application US/11114906
; Publication No. US20050266452A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0281-USA
; CURRENT APPLICATION NUMBER: US/11/114,906
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: US/10/413,437
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: US/10/020,079
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,103
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/289,422
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 764
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(764)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-11-114-906-6

Alignment Scores:

Pred. No.:	34.8	Length:	764
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0

DB: 7 Gaps: 0

US-10-677-865A-2 (1-1590) x US-11-114-906-6 (1-764)

```
Qy        253 GGTGGTGGCTCCAGTGGTAGCTCC 276
              ||||||||||||||||||||||||
Db        702 GlyGlyGlySerSerGlySerSer 709
```

RESULT 75
US-11-087-099-8165
; Sequence 8165, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8165
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-087-099-8165

Alignment Scores:

Pred. No.:	34.4	Length:	822
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	7	Gaps:	0

US-10-677-865A-2 (1-1590) x US-11-087-099-8165 (1-822)

```
Qy       1294 AGCAGCAGCTCCAGTTCCCAATCG 1317
              ||||||||||||||||||||||||
Db        381 SerSerSerSerSerSerGlnSer 388
```

Search completed: March 21, 2006, 22:33:08
Job time : 37 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2006, 22:22:59 ; Search time 6.7 Seconds
(without alignments)
3924.010 Million cell updates/sec

Title: US-10-677-865A-2
Perfect score: 528
Sequence: 1 atgggctcgtctcgggcacc.....agttactcgacagtccataa 1590

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1006954

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/absa/ABSSWEB_spool/US10677865/runat_21032006_152108_26891/app_query.fasta_1
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=oligo.rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=50 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absa02h
-USER=US10677865_@CGN_1_1_33_@runat_21032006_152108_26891 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	65.5	529	2	US-09-381-656-1	Sequence 1, Appli
2	123	23.3	529	2	US-10-164-230-61	Sequence 61, Appl
3	18	3.4	78	2	US-09-641-931F-39	Sequence 39, Appl
4	11	2.1	344	2	US-09-740-288A-8	Sequence 8, Appli
5	11	2.1	344	2	US-09-740-288A-20	Sequence 20, Appl
6	11	2.1	377	2	US-09-740-288A-22	Sequence 22, Appl
7	11	2.1	377	2	US-09-740-288A-24	Sequence 24, Appl
8	11	2.1	510	2	US-09-270-767-43633	Sequence 43633, A
9	10	1.9	365	1	US-08-515-251A-2	Sequence 2, Appli
10	10	1.9	406	2	US-10-104-047-3291	Sequence 3291, Ap
11	10	1.9	593	2	US-09-964-899-15	Sequence 15, Appl
12	10	1.9	645	2	US-09-964-899-17	Sequence 17, Appl
13	10	1.9	646	2	US-09-964-899-53	Sequence 53, Appl
14	9	1.7	13	1	US-08-257-341-1	Sequence 1, Appli
15	9	1.7	15	1	US-08-133-804-8	Sequence 8, Appli
16	9	1.7	15	1	US-08-461-838-8	Sequence 8, Appli
17	9	1.7	15	1	US-08-461-386-8	Sequence 8, Appli
18	9	1.7	15	1	US-08-356-786-12	Sequence 12, Appl
19	9	1.7	15	1	US-08-665-202-3	Sequence 3, Appli
20	9	1.7	15	2	US-08-483-749A-31	Sequence 31, Appl
21	9	1.7	15	2	US-09-315-574-3	Sequence 3, Appli
22	9	1.7	15	2	US-09-324-258-18	Sequence 18, Appl
23	9	1.7	17	2	US-09-333-213-4	Sequence 4, Appli
24	9	1.7	17	2	US-10-077-210-4	Sequence 4, Appli
25	9	1.7	31	2	US-09-324-258-19	Sequence 19, Appl
26	9	1.7	100	2	US-09-411-067C-5	Sequence 5, Appli
c 27	9	1.7	170	2	US-09-252-991A-26152	Sequence 26152, A
28	9	1.7	214	2	US-09-949-016-9109	Sequence 9109, Ap
29	9	1.7	243	1	US-08-133-804-6	Sequence 6, Appli
30	9	1.7	243	1	US-08-461-838-6	Sequence 6, Appli
31	9	1.7	243	1	US-08-461-386-6	Sequence 6, Appli
32	9	1.7	243	1	US-08-356-786-4	Sequence 4, Appli
33	9	1.7	246	1	US-08-257-341-7	Sequence 7, Appli
34	9	1.7	250	1	US-08-133-804-2	Sequence 2, Appli
35	9	1.7	250	1	US-08-461-184-8	Sequence 8, Appli
36	9	1.7	250	1	US-08-463-675-8	Sequence 8, Appli
37	9	1.7	250	1	US-08-464-589-8	Sequence 8, Appli
38	9	1.7	250	1	US-08-461-838-2	Sequence 2, Appli
39	9	1.7	250	1	US-08-461-386-2	Sequence 2, Appli
40	9	1.7	250	2	US-09-248-796A-22487	Sequence 22487, A
41	9	1.7	252	1	US-08-133-804-4	Sequence 4, Appli
42	9	1.7	252	1	US-08-461-838-4	Sequence 4, Appli
43	9	1.7	252	1	US-08-461-386-4	Sequence 4, Appli
44	9	1.7	271	2	US-09-248-796A-20590	Sequence 20590, A
45	9	1.7	367	1	US-08-257-341-5	Sequence 5, Appli
46	9	1.7	367	1	US-08-515-251A-4	Sequence 4, Appli
47	9	1.7	440	2	US-10-012-231A-52	Sequence 52, Appl
48	9	1.7	440	2	US-10-015-389A-52	Sequence 52, Appl
49	9	1.7	440	2	US-10-006-768A-52	Sequence 52, Appl
50	9	1.7	440	2	US-10-015-671A-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Alignment Scores:

Pred. No.:	3.79e-295	Length:	529
Score:	346.00	Matches:	346
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	65.5%	Indels:	0

```
DB:                 2           Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-381-656-1 (1-529)

Qy       550 CGCGGAATACTAAACCCTTCCCAGCCTGGACAAAGCTCTTCCTCTTCCCAAACCTCTGGG 609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       184 ArgGlyIleLeuAsnProSerGlnProGlyGlnSerSerSerSerSerGlnThrSerGly 203

Qy       610 GTATCCAGCAGTGGCCAAAGCGTCAGCTCCAACCAGCGTCCCTGTAGTTCGGACATCCCC 669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       204 ValSerSerSerGlyGlnSerValSerSerAsnGlnArgProCysSerSerAspIlePro 223

Qy       670 GACTCTCCCTGCAGTGGAGGGCCCATCGTCTCGCACTCTGGCCCCTACATCCCCAGCTCC 729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       224 AspSerProCysSerGlyGlyProIleValSerHisSerGlyProTyrIleProSerSer 243

Qy       730 CACTCTGTGTCAGGGGGTCAGAGGCCTGTGGTGGTGGTGGTGGACCAGCACGGTTCTGGT 789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       244 HisSerValSerGlyGlyGlnArgProValValValValValAspGlnHisGlySerGly 263

Qy       790 GCCCCTGGAGTGGTTCAAGGTCCCCCCTGTAGCAATGGTGGCCTTCCAGGCAAGCCCTGT 849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       264 AlaProGlyValValGlnGlyProProCysSerAsnGlyGlyLeuProGlyLysProCys 283

Qy       850 CCCCCAATCACCTCTGTAGACAAATCCTATGGTGGCTACGAGGTGGTGGGTGGCTCCTCT 909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       284 ProProIleThrSerValAspLysSerTyrGlyGlyTyrGluValValGlyGlySerSer 303

Qy       910 GACAGTTATCTGGTTCCAGGCATGACCTACAGTAAGGGTAAAATCTATCCTGTGGGCTAC 969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       304 AspSerTyrLeuValProGlyMetThrTyrSerLysGlyLysIleTyrProValGlyTyr 323

Qy       970 TTCACCAAAGAGAACCCTGTGAAAGGCTCTCCAGGGGTCCCTTCCTTTGCAGCTGGGCCC 1029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       324 PheThrLysGluAsnProValLysGlySerProGlyValProSerPheAlaAlaGlyPro 343

Qy      1030 CCCATCTCTGAGGGCAAATACTTCTCCAGCAACCCCATCATCCCCAGCCAGTCGGCAGCT 1089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       344 ProIleSerGluGlyLysTyrPheSerSerAsnProIleIleProSerGlnSerAlaAla 363

Qy      1090 TCCTCGGCCATTGCGTTCCAGCCAGTGGGGACTGGTGGGGTCCAGCTCTGTGGAGGCGGC 1149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       364 SerSerAlaIleAlaPheGlnProValGlyThrGlyGlyValGlnLeuCysGlyGlyGly 383

Qy      1150 TCCACGGGCTCCAAGGGACCCTGCTCTCCCTCCAGTTCTCGAGTCCCCAGCAGTTCTAGC 1209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       384 SerThrGlySerLysGlyProCysSerProSerSerSerArgValProSerSerSerSer 403

Qy      1210 ATTTCCAGCAGCGCCGGTTCACCCTACCATCCCTGCGGCAGTGCTTCCCAGAGCCCCTGC 1269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       404 IleSerSerSerAlaGlySerProTyrHisProCysGlySerAlaSerGlnSerProCys 423

Qy      1270 TCCCCACCAGGCACCGGCTCCTTCAGCAGCAGCTCCAGTTCCCAATCGAGTGGCAAAATC 1329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       424 SerProProGlyThrGlySerPheSerSerSerSerSerGlnSerSerGlyLysIle 443

Qy      1330 ATCCTTCAGCCTTGTGGCAGCAAGTCCAGCTCTTCTGGTCACCCTTGCATGTCTGTCTCC 1389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       444 IleLeuGlnProCysGlySerLysSerSerSerSerGlyHisProCysMetSerValSer 463

Qy      1390 TCCTTGACACTGACTGGGGGCCCCGATGGCTCTCCCCATCCTGATCCCTCCGCTGGTGCC 1449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       464 SerLeuThrLeuThrGlyGlyProAspGlySerProHisProAspProSerAlaGlyAla 483

Qy      1450 AAGCCCTGTGGCTCCAGCAGTGCTGGAAAGATCCCCTGCCGCTCCATCCGGGATATCCTA 1509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       484 LysProCysGlySerSerSerAlaGlyLysIleProCysArgSerIleArgAspIleLeu 503

Qy      1510 GCCCAAGTGAAGCCTCTGGGGCCCCAGCTAGCTGACCCTGAAGTTTTCCTACCCCAAGGA 1569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       504 AlaGlnValLysProLeuGlyProGlnLeuAlaAspProGluValPheLeuProGlnGly 523

Qy      1570 GAGTTACTCGACAGTCCA 1587
```

```
             ||||||||||||||||||
Db       524 GluLeuLeuAspSerPro 529

RESULT 2
US-10-164-230-61
; Sequence 61, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

Alignment Scores:
Pred. No.:              2.01e-99        Length:         529
Score:                  123.00          Matches:        123
```

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            23.3%           Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-164-230-61 (1-529)

Qy          1 ATGGGCTCGTCTCGGGCACCCTGGATGGGGCGTGTGGGTGGGCACGGGATGATGGCACTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerSerArgAlaProTrpMetGlyArgValGlyGlyHisGlyMetMetAlaLeu 20

Qy         61 CTGCTGGCTGGTCTCCTCCTGCCAGGGACCTTGGCTAAGAGCATTGGCACCTTCTCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuLeuAlaGlyLeuLeuLeuProGlyThrLeuAlaLysSerIleGlyThrPheSerAsp 40

Qy        121 CCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTCACTGGGAAGGGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProCysLysAspProThrArgIleThrSerProAsnAspProCysLeuThrGlyLysGly 60

Qy        181 GACTCCAGCGGCTTCAGTAGCTACAGTGGCTCCAGCAGTTCTGGCAGCTCCATTTCCAGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AspSerSerGlyPheSerSerTyrSerGlySerSerSerSerGlySerSerIleSerSer 80

Qy        241 GCCAGAAGCTCTGGTGGTGGCTCCAGTGGTAGCTCCAGCGGATCCAGCATTGCCCAGGGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaArgSerSerGlyGlyGlySerSerGlySerSerSerSerGlySerSerIleAlaGlnGly 100

Qy        301 GGTTCTGCAGGATCTTTTAAGCCAGGAACGGGGTATTCCCAGGTCAGCTACTCCTCCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlySerAlaGlySerPheLysProGlyThrGlyTyrSerGlnValSerTyrSerSerGly 120

Qy        361 TCTGGCTCT 369
              |||||||||
Db        121 SerGlySer 123

RESULT 3
US-09-641-931F-39
; Sequence 39, Application US/09641931F
; Patent No. 6830885
; GENERAL INFORMATION:
;  APPLICANT: Lanctot, et al.
;  TITLE OF INVENTION: Nucleic Acid Molecule, Method and Kit for Selecting a
;  TITLE OF INVENTION:  Nucleic Acid Having A Desired Feature
;  FILE REFERENCE: 2003390-0001
;  CURRENT APPLICATION NUMBER: US/09/641,931F
;  CURRENT FILING DATE:  2003-04-08
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-641-931F-39

Alignment Scores:
Pred. No.:              3.5e-07         Length:        78
Score:                  18.00           Matches:       18
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            3.4%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-641-931F-39 (1-78)

Qy        115 TCAGACCCCTGTAAGGACCCCACGCGTATCACCTCCCCTAACGACCCCTGCCTC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 SerAspProCysLysAspProThrArgIleThrSerProAsnAspProCysLeu 55

RESULT 4
US-09-740-288A-8
; Sequence 8, Application US/09740288A
; Patent No. 6849783
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  APPLICANT:  Kinney, Anthony
;  APPLICANT:  Miao, Guo-Hua
;  APPLICANT:  Orozco, Emil
;  TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
;  FILE REFERENCE: BB1429 US NA
;  CURRENT APPLICATION NUMBER: US/09/740,288A
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER: US 60/172929
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-740-288A-8

Alignment Scores:
Pred. No.:              0.418           Length:        344
Score:                  11.00           Matches:       11
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            2.1%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-740-288A-8 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 5
US-09-740-288A-20
; Sequence 20, Application US/09740288A
; Patent No. 6849783
; GENERAL INFORMATION:
;  APPLICANT: Allen, Stephen
;  APPLICANT:  Kinney, Anthony
;  APPLICANT:  Miao, Guo-Hua
;  APPLICANT:  Orozco, Emil
;  TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
;  FILE REFERENCE: BB1429 US NA
;  CURRENT APPLICATION NUMBER: US/09/740,288A
;  CURRENT FILING DATE:  2000-12-19
;  PRIOR APPLICATION NUMBER: US 60/172929
;  PRIOR FILING DATE: 1999-12-21
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-740-288A-20

Alignment Scores:
Pred. No.:              0.418           Length:        344
Score:                  11.00           Matches:       11
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            2.1%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-740-288A-20 (1-344)

Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29

RESULT 6
US-09-740-288A-22
; Sequence 22, Application US/09740288A
; Patent No. 6849783
```

; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-09-740-288A-22

Alignment Scores:

Pred. No.:	0.415	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-740-288A-22 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 7
US-09-740-288A-24
; Sequence 24, Application US/09740288A
; Patent No. 6849783
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
US-09-740-288A-24

Alignment Scores:

Pred. No.:	0.415	Length:	377
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-740-288A-24 (1-377)

```
Qy        467 GCAGCAGCAGCAGCTTTCAGTTCAGCAGCAGCA 499
              |||||||||||||||||||||||||||||||||
Db         19 AlaAlaAlaAlaAlaPheSerSerAlaAlaAla 29
```

RESULT 8
US-09-270-767-43633
; Sequence 43633, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43633
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43633

Alignment Scores:

Pred. No.:	0.402	Length:	510
Score:	11.00	Matches:	11
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	2.1%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-270-767-43633 (1-510)

```
Qy        448 AGCAGCAGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||||||
Db        478 SerSerSerSerHisSerSerSerSerSerSer 488
```

RESULT 9
US-08-515-251A-2
; Sequence 2, Application US/08515251A
; Patent No. 5891677
; GENERAL INFORMATION:
; APPLICANT: GERLACH, GERALD F.
; APPLICANT: WILLSON, PHILIP J.
; APPLICANT: ROSSI-CAMPOS, AMALIA
; APPLICANT: POTTER, ANDREW A.
; TITLE OF INVENTION: ACTINOBACILLUS PLEUROPNEUMONIAE OUTER
; TITLE OF INVENTION: MEMBRANE LIPOPROTEIN A AND USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/515,251A
; FILING DATE: 15-AUG-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,558
; FILING DATE: 05-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0027.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-515-251A-2

Alignment Scores:

Pred. No.:	3.14	Length:	365
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	1	Gaps:	0

US-10-677-865A-2 (1-1590) x US-08-515-251A-2 (1-365)

```
Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSerSer 30
```

RESULT 10
US-10-104-047-3291
; Sequence 3291, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3291
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3291

Alignment Scores:

Pred. No.:	3.1	Length:	406
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-104-047-3291 (1-406)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db         72 GlyGlySerSerGlySerSerSerGlySer 81
```

RESULT 11
US-09-964-899-15
; Sequence 15, Application US/09964899
; Patent No. 6900367
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-15

Alignment Scores:

Pred. No.:	2.99	Length:	593
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-964-899-15 (1-593)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        261 GlyGlySerSerGlySerSerSerGlySer 270
```

RESULT 12
US-09-964-899-17
; Sequence 17, Application US/09964899
; Patent No. 6900367
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-17

Alignment Scores:

Pred. No.:	2.96	Length:	645
Score:	10.00	Matches:	10
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.9%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-964-899-17 (1-645)

```
Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        311 GlyGlySerSerGlySerSerSerGlySer 320
```

RESULT 13
US-09-964-899-53
; Sequence 53, Application US/09964899
; Patent No. 6900367
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 646
; TYPE: PRT

```
;   ORGANISM: Homo sapiens
US-09-964-899-53

Alignment Scores:
Pred. No.:              2.96          Length:          646
Score:                  10.00         Matches:         10
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.9%          Indels:          0
DB:                     2             Gaps:            0

US-10-677-865A-2 (1-1590) x US-09-964-899-53 (1-646)

Qy        256 GGTGGCTCCAGTGGTAGCTCCAGCGGATCC 285
              ||||||||||||||||||||||||||||||
Db        312 GlyGlySerSerGlySerSerSerGlySer 321

RESULT 14
US-08-257-341-1
; Sequence 1, Application US/08257341
; Patent No. 5525491
;  GENERAL INFORMATION:
;    APPLICANT:  HUSTON, JAMES S
;    APPLICANT:  OPPERMANN, HERMANN
;    APPLICANT:  TIMASHEFF, SERGE N
;    TITLE OF INVENTION:  SERINE RICH PEPTIDE LINKER
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CREATIVE BIOMOLECULES, INC./PATENT DEPT.
;      STREET:  35 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/257,341
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/842,149
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/662,226
;      FILING DATE:  27-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL ESQ, PAULA A
;      REGISTRATION NUMBER:  32,503
;      REFERENCE/DOCKET NUMBER:  CRP-064CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000 (ATTY)
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..13
;      OTHER INFORMATION:  /note= "(SER)4-GLY LINKER.  THE
;      OTHER INFORMATION:  REPEATING SEQUENCE  "(SER)4-GLY" (E.G., RES. 3-7)
;      OTHER INFORMATION:  MAY BE REPEATED MULTIPLE TIMES (SEE SPECIFICATION.)
US-08-257-341-1

Alignment Scores:
Pred. No.:              33.3          Length:          13
Score:                  9.00          Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.7%          Indels:          0
DB:                     1             Gaps:            0

US-10-677-865A-2 (1-1590) x US-08-257-341-1 (1-13)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          1 SerGlySerSerSerSerGlySerSer 9

RESULT 15
US-08-133-804-8
; Sequence 8, Application US/08133804
; Patent No. 5534254
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/133,804
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..15
;      OTHER INFORMATION:  /note= "LINKER 2"
US-08-133-804-8

Alignment Scores:
Pred. No.:              32.8          Length:          15
Score:                  9.00          Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            1.7%          Indels:          0
DB:                     1             Gaps:            0

US-10-677-865A-2 (1-1590) x US-08-133-804-8 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 16
```

```
US-08-461-838-8
; Sequence 8, Application US/08461838
; Patent No. 5753204
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,838
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..15
;      OTHER INFORMATION:  /note= "LINKER 2"
US-08-461-838-8

Alignment Scores:
Pred. No.:                32.8          Length:         15
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-838-8 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 17
US-08-461-386-8
; Sequence 8, Application US/08461386
; Patent No. 5837846
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,386
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..15
;      OTHER INFORMATION:  /note= "LINKER 2"
US-08-461-386-8

Alignment Scores:
Pred. No.:                32.8          Length:         15
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-386-8 (1-15)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 18
US-08-356-786-12
; Sequence 12, Application US/08356786
; Patent No. 5877305
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Protein for Cancer
;    TITLE OF INVENTION:  Marker
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edmund R. Pitcher, Testa, Hurwitz, & Thibeault
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/356,786
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/831,967
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;      REGISTRATION NUMBER:  27,829
;      REFERENCE/DOCKET NUMBER:  CRP-053
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-356-786-12

Alignment Scores:
Pred. No.:                32.8          Length:         15
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-356-786-12 (1-15)

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
               |||||||||||||||||||||||||||
Db           4 SerGlySerSerSerSerGlySerSer 12

RESULT 19
US-08-665-202-3
; Sequence 3, Application US/08665202
; Patent No. 5977322
;  GENERAL INFORMATION:
;    APPLICANT:  Marks, James D.
;    APPLICANT:  Schier, Robert
;    TITLE OF INVENTION:  No. 5977322el High Affinity Human Antibodies to
;    TITLE OF INVENTION:  Tumor Antigens
;    NUMBER OF SEQUENCES:  141
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,202
;      FILING DATE:  13-JUN-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,238
;      FILING DATE:  14-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,250
;      FILING DATE:  15-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  02307E-061410
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-665-202-3

Alignment Scores:
Pred. No.:                32.8          Length:         15
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-665-202-3 (1-15)

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
               |||||||||||||||||||||||||||
Db           4 SerGlySerSerSerSerGlySerSer 12

RESULT 20
US-08-483-749A-31
; Sequence 31, Application US/08483749A
; Patent No. 6054561
;  GENERAL INFORMATION:
;    APPLICANT:  RING, DAVID B.
;    TITLE OF INVENTION:  ANTIGEN-BINDING SITES OF ANTIBODY
;    TITLE OF INVENTION:  MOLECULES SPECIFIC FOR CANCER ANTIGENS
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHIRON CORPORATION
;      STREET:  INTELLECTUAL PROPERTY - R440, PO BOX 8097
;      CITY:  EMERYVILLE
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94662-8097
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,749A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SAVEREIDE, PAUL B.
;      REGISTRATION NUMBER:  36,914
;      REFERENCE/DOCKET NUMBER:  0508.008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 601-2585
;      TELEFAX:  (510) 655-3542
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-483-749A-31

Alignment Scores:
Pred. No.:                32.8          Length:         15
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       2             Gaps:           0
```

US-10-677-865A-2 (1-1590) x US-08-483-749A-31 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 21
US-09-315-574-3
; Sequence 3, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-315-574-3

Alignment Scores:

Pred. No.:	32.8	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-315-574-3 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 22
US-09-324-258-18
; Sequence 18, Application US/09324258
; Patent No. 6723517
; GENERAL INFORMATION:
; APPLICANT: Bamdad, Cynthia C.
; TITLE OF INVENTION: THE USE OF SELF-ASSEMBLED MONOLAYERS TO
; TITLE OF INVENTION: PROBE THE STRUCTURE OF A TARGET MOLECULE
; FILE REFERENCE: M1015/7004/TJO
; CURRENT APPLICATION NUMBER: US/09/324,258
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: U.S. 60/087,766
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker
US-09-324-258-18

Alignment Scores:

Pred. No.:	32.8	Length:	15
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-324-258-18 (1-15)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 23
US-09-333-213-4
; Sequence 4, Application US/09333213
; Patent No. 6548653
; GENERAL INFORMATION:
; APPLICANT: Young, Michael
; APPLICANT: Meade, Harry
; APPLICANT: Krane, Ian
; TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
; FILE REFERENCE: 10275/041001
; CURRENT APPLICATION NUMBER: US/09/333,213
; CURRENT FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated linker sequence
US-09-333-213-4

Alignment Scores:

Pred. No.:	32.4	Length:	17
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-333-213-4 (1-17)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12
```

RESULT 24

```
US-10-077-210-4
; Sequence 4, Application US/10077210
; Patent No. 6800462
; GENERAL INFORMATION:
;  APPLICANT: Wu, Chung-Hsiun
;  APPLICANT:  Lin, Rong-Hwa
;  APPLICANT:  Hsu, Pei-Ling
;  TITLE OF INVENTION: PRODUCTION OF RECOMBINANT PROTEINS IN VIVO AND USE FOR GENERATING
;  TITLE OF INVENTION:  ANTIBODIES
;  FILE REFERENCE: 13062-004001
;  CURRENT APPLICATION NUMBER: US/10/077,210
;  CURRENT FILING DATE:  2002-06-13
;  PRIOR APPLICATION NUMBER: 60/318,474
;  PRIOR FILING DATE: 2001-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 17
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Linker sequence
US-10-077-210-4

Alignment Scores:
Pred. No.:              32.4          Length:         17
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-077-210-4 (1-17)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 25
US-09-324-258-19
; Sequence 19, Application US/09324258
; Patent No. 6723517
; GENERAL INFORMATION:
;  APPLICANT:  Bamdad, Cynthia C.
;  TITLE OF INVENTION:  THE USE OF SELF-ASSEMBLED MONOLAYERS TO
;  TITLE OF INVENTION:  PROBE THE STRUCTURE OF A TARGET MOLECULE
;  FILE REFERENCE:  M1015/7004/TJO
;  CURRENT APPLICATION NUMBER: US/09/324,258
;  CURRENT FILING DATE:  1999-06-02
;  PRIOR APPLICATION NUMBER:  U.S. 60/087,766
;  PRIOR FILING DATE:  1998-06-02
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 1X-linker-1X
US-09-324-258-19

Alignment Scores:
Pred. No.:              30.5          Length:         31
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-324-258-19 (1-31)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db         12 SerGlySerSerSerSerGlySerSer 20

RESULT 26
US-09-411-067C-5
; Sequence 5, Application US/09411067C
; Patent No. 6576610
; GENERAL INFORMATION:
;  APPLICANT: NUVAS, LLC
;  APPLICANT:  HOUSTON, L.L.
;  TITLE OF INVENTION: USE OF A CONTEXT-DEPENDENT FUNCTIONAL ENTITY TO ENHANCE THE EFFI
;  TITLE OF INVENTION:  AN AGENT
;  FILE REFERENCE: NUVAS1140
;  CURRENT APPLICATION NUMBER: US/09/411,067C
;  CURRENT FILING DATE:  2002-05-07
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Spacer element
;   FEATURE:
;   NAME/KEY: REPEAT
;   LOCATION: (1)..(5)
;   OTHER INFORMATION: Spacer peptide variation; sequence repeated "n" times, where n =
;   OTHER INFORMATION: 1-20
US-09-411-067C-5

Alignment Scores:
Pred. No.:              27            Length:         100
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-411-067C-5 (1-100)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db          4 SerGlySerSerSerSerGlySerSer 12

RESULT 27
US-09-252-991A-26152
; Sequence 26152, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26152
;   LENGTH: 170
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26152

Alignment Scores:
Pred. No.:              25.6          Length:         170
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0
```

US-10-677-865A-2 (1-1590) x US-09-252-991A-26152 (1-170)

```
Qy        426 TCCCGAGTGAGAGCTGCTGCTCCCCAG 400
              ||||||||||||||||||||||||||||
Db        135 SerArgValArgAlaAlaAlaProGln 143
```

RESULT 28
US-09-949-016-9109
; Sequence 9109, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9109
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9109

Alignment Scores:

Pred. No.:	25	Length:	214
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-949-016-9109 (1-214)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        119 SerSerGlySerSerSerGlySerSer 127
```

RESULT 29
US-08-133-804-6
; Sequence 6, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibeault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-133-804-6

Alignment Scores:

Pred. No.:	24.7	Length:	243
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	1	Gaps:	0

US-10-677-865A-2 (1-1590) x US-08-133-804-6 (1-243)

```
Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129
```

RESULT 30
US-08-461-838-6
; Sequence 6, Application US/08461838
; Patent No. 5753204
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibeault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,838
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-838-6

Alignment Scores:

```
Pred. No.:              24.7          Length:         243
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-838-6 (1-243)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 31
US-08-461-386-6
; Sequence 6, Application US/08461386
; Patent No. 5837846
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,386
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  243 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-461-386-6

Alignment Scores:
Pred. No.:              24.7          Length:         243
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-386-6 (1-243)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 32
US-08-356-786-4
; Sequence 4, Application US/08356786
; Patent No. 5877305
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Protein for Cancer
;    TITLE OF INVENTION:  Marker
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edmund R. Pitcher, Testa, Hurwitz, & Thibeault
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/356,786
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/831,967
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;      REGISTRATION NUMBER:  27,829
;      REFERENCE/DOCKET NUMBER:  CRP-053
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  243 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-356-786-4

Alignment Scores:
Pred. No.:              24.7          Length:         243
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-356-786-4 (1-243)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 33
US-08-257-341-7
; Sequence 7, Application US/08257341
; Patent No. 5525491
;  GENERAL INFORMATION:
;    APPLICANT:  HUSTON, JAMES S
;    APPLICANT:  OPPERMANN, HERMANN
;    APPLICANT:  TIMASHEFF, SERGE N
;    TITLE OF INVENTION:  SERINE RICH PEPTIDE LINKER
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CREATIVE BIOMOLECULES, INC./PATENT DEPT.
;      STREET:  35 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
```

```
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/257,341
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/842,149
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/662,226
;      FILING DATE:  27-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL ESQ, PAULA A
;      REGISTRATION NUMBER:  32,503
;      REFERENCE/DOCKET NUMBER:  CRP-064CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000 (ATTY)
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-257-341-7

Alignment Scores:
Pred. No.:                24.7          Length:         246
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-257-341-7 (1-246)

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
               ||||||||||||||||||||||||||
Db         121 SerGlySerSerSerSerGlySerSer 129

RESULT 34
US-08-133-804-2
; Sequence 2, Application US/08133804
; Patent No. 5534254
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/133,804
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-133-804-2

Alignment Scores:
Pred. No.:                24.6          Length:         250
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-133-804-2 (1-250)

Qy         205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
               ||||||||||||||||||||||||||
Db         125 SerGlySerSerSerSerGlySerSer 133

RESULT 35
US-08-461-184-8
; Sequence 8, Application US/08461184
; Patent No. 5631158
;  GENERAL INFORMATION:
;    APPLICANT:  DORAI, HAIMANTI
;    APPLICANT:  OPPERMANN, HERMANN
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR HIGH PROTEIN
;    TITLE OF INVENTION:  PRODUCTION FROM NON-NATIVE DNA
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
;      STREET:  45 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  07148
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,184
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/143,498
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KELLEY, ROBIN D
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  CRP093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000
;      TELEFAX:  617/248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-461-184-8

Alignment Scores:
Pred. No.:                24.6          Length:         250
```

```
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-184-8 (1-250)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 36
US-08-463-675-8
; Sequence 8, Application US/08463675
; Patent No. 5658763
;  GENERAL INFORMATION:
;    APPLICANT:  DORAI, HAIMANTI
;    APPLICANT:  OPPERMANN, HERMANN
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR HIGH PROTEIN
;    TITLE OF INVENTION:  PRODUCTION FROM NON-NATIVE DNA
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
;      STREET:  45 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  07148
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,675
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/143,498
;      FILING DATE:  25-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KELLEY, ROBIN D
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  CRP093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000
;      TELEFAX:  617/248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-463-675-8

Alignment Scores:
Pred. No.:              24.6          Length:         250
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-463-675-8 (1-250)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 37
US-08-464-589-8
; Sequence 8, Application US/08464589
; Patent No. 5733782
;  GENERAL INFORMATION:
;    APPLICANT:  DORAI, HAIMANTI
;    APPLICANT:  OPPERMANN, HERMANN
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR HIGH PROTEIN
;    TITLE OF INVENTION:  PRODUCTION FROM NON-NATIVE DNA
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
;      STREET:  45 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  07148
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,589
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/143,498
;      FILING DATE:  25-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KELLEY, ROBIN D
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  CRP093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000
;      TELEFAX:  617/248-7100
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-464-589-8

Alignment Scores:
Pred. No.:              24.6          Length:         250
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-464-589-8 (1-250)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 38
US-08-461-838-2
; Sequence 2, Application US/08461838
; Patent No. 5753204
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
```

```
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/461,838
;       FILING DATE:
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kelley, Robin D.
;       REGISTRATION NUMBER:  34,637
;       REFERENCE/DOCKET NUMBER:  2054/22
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-248-7477
;       TELEFAX:  617-248-7100
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  250 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-461-838-2

Alignment Scores:
Pred. No.:                24.6          Length:         250
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-838-2 (1-250)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 39
US-08-461-386-2
; Sequence 2, Application US/08461386
; Patent No. 5837846
;  GENERAL INFORMATION:
;     APPLICANT:  Huston, James S.
;     APPLICANT:  Oppermann, Hermann
;     APPLICANT:  Houston, L. L.
;     APPLICANT:  Ring, David B.
;     TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;     TITLE OF INVENTION:  Imaging
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;       STREET:  Exchange Place, 53 State Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/461,386
;       FILING DATE:
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kelley, Robin D.
;       REGISTRATION NUMBER:  34,637
;       REFERENCE/DOCKET NUMBER:  2054/22
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-248-7477
;       TELEFAX:  617-248-7100
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  250 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-461-386-2

Alignment Scores:
Pred. No.:                24.6          Length:         250
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-461-386-2 (1-250)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              |||||||||||||||||||||||||||
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 40
US-09-248-796A-22487
; Sequence 22487, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22487
;   LENGTH: 250
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-22487

Alignment Scores:
Pred. No.:                24.6          Length:         250
Score:                    9.00          Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              1.7%          Indels:         0
DB:                       2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-248-796A-22487 (1-250)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              |||||||||||||||||||||||||||
Db         71 SerSerGlySerSerSerGlySerSer 79

RESULT 41
US-08-133-804-4
; Sequence 4, Application US/08133804
; Patent No. 5534254
;  GENERAL INFORMATION:
;     APPLICANT:  Huston, James S.
;     APPLICANT:  Oppermann, Hermann
;     APPLICANT:  Houston, L. L.
;     APPLICANT:  Ring, David B.
;     TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;     TITLE OF INVENTION:  Imaging
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
```

```
;       STREET:  Exchange Place, 53 State Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/133,804
;       FILING DATE:
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kelley, Robin D.
;       REGISTRATION NUMBER:  34,637
;       REFERENCE/DOCKET NUMBER:  2054/22
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-248-7477
;       TELEFAX:  617-248-7100
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  252 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-133-804-4

Alignment Scores:
Pred. No.:                24.6          Length:          252
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       1             Gaps:            0

US-10-677-865A-2 (1-1590) x US-08-133-804-4 (1-252)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 42
US-08-461-838-4
; Sequence 4, Application US/08461838
; Patent No. 5753204
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,838
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-461-838-4

Alignment Scores:
Pred. No.:                24.6          Length:          252
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       1             Gaps:            0

US-10-677-865A-2 (1-1590) x US-08-461-838-4 (1-252)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 43
US-08-461-386-4
; Sequence 4, Application US/08461386
; Patent No. 5837846
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Proteins For
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Testa, Hurwitz & Thibeault/Patent Department
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,386
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelley, Robin D.
;      REGISTRATION NUMBER:  34,637
;      REFERENCE/DOCKET NUMBER:  2054/22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7477
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-461-386-4

Alignment Scores:
Pred. No.:                24.6          Length:          252
Score:                    9.00          Matches:         9
```

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     1               Gaps:          0

US-10-677-865A-2 (1-1590) x US-08-461-386-4 (1-252)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        121 SerGlySerSerSerSerGlySerSer 129

RESULT 44
US-09-248-796A-20590
; Sequence 20590, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20590
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (44)
;   OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkn
US-09-248-796A-20590

Alignment Scores:
Pred. No.:              24.4            Length:        271
Score:                  9.00            Matches:       9
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     2               Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-248-796A-20590 (1-271)

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCCAGC 279
              ||||||||||||||||||||||||||
Db        194 GlyGlyGlySerSerGlySerSerSer 202

RESULT 45
US-08-257-341-5
; Sequence 5, Application US/08257341
; Patent No. 5525491
;  GENERAL INFORMATION:
;    APPLICANT:  HUSTON, JAMES S
;    APPLICANT:  OPPERMANN, HERMANN
;    APPLICANT:  TIMASHEFF, SERGE N
;    TITLE OF INVENTION:  SERINE RICH PEPTIDE LINKER
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CREATIVE BIOMOLECULES, INC./PATENT DEPT.
;      STREET:  35 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/257,341
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/842,149
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/662,226
;      FILING DATE:  27-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL ESQ, PAULA A
;      REGISTRATION NUMBER:  32,503
;      REFERENCE/DOCKET NUMBER:  CRP-064CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000 (ATTY)
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  367 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-257-341-5

Alignment Scores:
Pred. No.:              23.7            Length:        367
Score:                  9.00            Matches:       9
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            1.7%            Indels:        0
DB:                     1               Gaps:          0

US-10-677-865A-2 (1-1590) x US-08-257-341-5 (1-367)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        242 SerGlySerSerSerSerGlySerSer 250

RESULT 46
US-08-515-251A-4
; Sequence 4, Application US/08515251A
; Patent No. 5891677
;  GENERAL INFORMATION:
;    APPLICANT:  GERLACH, GERALD F.
;    APPLICANT:  WILLSON, PHILIP J.
;    APPLICANT:  ROSSI-CAMPOS, AMALIA
;    APPLICANT:  POTTER, ANDREW A.
;    TITLE OF INVENTION:  ACTINOBACILLUS PLEUROPNEUMONIAE OUTER
;    TITLE OF INVENTION:  MEMBRANE LIPOPROTEIN A AND USES THEREOF
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBINS & ASSOCIATES
;      STREET:  90 MIDDLEFIELD ROAD, SUITE 200
;      CITY:  MENLO PARK
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/515,251A
;      FILING DATE:  15-AUG-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/971,558
;      FILING DATE:  05-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0027.10
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (650) 325-7812
;      TELEFAX:  (650) 325-7823
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  367 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-515-251A-4

Alignment Scores:
Pred. No.:              23.7          Length:         367
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     1             Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-515-251A-4 (1-367)

Qy        250 TCTGGTGGTGGCTCCAGTGGTAGCTCC 276
              |||||||||||||||||||||||||||
Db         21 SerGlyGlyGlySerSerGlySerSer 29

RESULT 47
US-10-012-231A-52
; Sequence 52, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan 1.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C23
;  CURRENT APPLICATION NUMBER: US/10/012,231A
;  CURRENT FILING DATE:  2002-06-10
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-231A-52

Alignment Scores:
Pred. No.:              23.2          Length:         440
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-012-231A-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 48
US-10-015-389A-52
; Sequence 52, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan 1.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C48
;  CURRENT APPLICATION NUMBER: US/10/015,389A
;  CURRENT FILING DATE:  2002-06-25
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-015-389A-52

Alignment Scores:
Pred. No.:              23.2          Length:         440
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-015-389A-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 49
US-10-006-768A-52
; Sequence 52, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan 1.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C10
;  CURRENT APPLICATION NUMBER: US/10/006,768A
;  CURRENT FILING DATE:  2002-03-05
;  NUMBER OF SEQ ID NOS: 477
;  Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

US-10-006-768A-52

Alignment Scores:

Pred. No.:	23.2	Length:	440
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-006-768A-52 (1-440)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281
```

RESULT 50
US-10-015-671A-52
; Sequence 52, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; CURRENT FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-52

Alignment Scores:

Pred. No.:	23.2	Length:	440
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-015-671A-52 (1-440)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281
```

RESULT 51
US-10-015-393A-52
; Sequence 52, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-393A-52

Alignment Scores:

Pred. No.:	23.2	Length:	440
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-015-393A-52 (1-440)

```
Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281
```

RESULT 52
US-10-011-833A-52
; Sequence 52, Application US/10011833A
; Patent No. 6951920
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan l.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C22
; CURRENT APPLICATION NUMBER: US/10/011,833A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-833A-52

Alignment Scores:

Pred. No.:	23.2	Length:	440
Score:	9.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.7%	Indels:	0

```
DB:                    2              Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-011-833A-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 53
US-10-006-041A-52
; Sequence 52, Application US/10006041A
; Patent No. 6951921
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan l.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C8
;  CURRENT APPLICATION NUMBER: US/10/006,041A
;  CURRENT FILING DATE:  2001-12-06
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-006-041A-52

Alignment Scores:
Pred. No.:              23.2          Length:         440
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-006-041A-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 54
US-10-012-064A-52
; Sequence 52, Application US/10012064A
; Patent No. 6953841
; GENERAL INFORMATION:
;  APPLICANT: Baker, Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan l.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2830P1C19
;  CURRENT APPLICATION NUMBER: US/10/012,064A
;  CURRENT FILING DATE:  2002-07-15
;  PRIOR APPLICATION NUMBER: 60/098716
;  PRIOR FILING DATE: 1998-09-01
;  PRIOR APPLICATION NUMBER: 60/098723
;  PRIOR FILING DATE: 1998-09-01
;  PRIOR APPLICATION NUMBER: 60/098749
;  PRIOR FILING DATE: 1998-09-01
;  PRIOR APPLICATION NUMBER: 60/098750
;  PRIOR FILING DATE: 1998-09-01
;  PRIOR APPLICATION NUMBER: 60/098803
;  PRIOR FILING DATE: 1998-09-02
;  PRIOR APPLICATION NUMBER: 60/098821
;  PRIOR FILING DATE: 1998-09-02
;  PRIOR APPLICATION NUMBER: 60/098843
;  PRIOR FILING DATE: 1998-09-02
;  PRIOR APPLICATION NUMBER: 60/099536
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 60/099596
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 60/099598
;  PRIOR FILING DATE: 1998-09-09
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 52
;   LENGTH: 440
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-064A-52

Alignment Scores:
Pred. No.:              23.2          Length:         440
Score:                  9.00          Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            1.7%          Indels:         0
DB:                     2             Gaps:           0

US-10-677-865A-2 (1-1590) x US-10-012-064A-52 (1-440)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        273 SerSerGlySerSerSerGlySerSer 281

RESULT 55
US-09-949-016-6315
; Sequence 6315, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6315
;   LENGTH: 441
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-6315

Alignment Scores:
```

```
Pred. No.:              23.2          Length:        441
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-949-016-6315 (1-441)

Qy        262 TCCAGTGGTAGCTCCAGCGGATCCAGC 288
              ||||||||||||||||||||||||||||
Db        346 SerSerGlySerSerSerGlySerSer 354

RESULT 56
US-08-257-341-9
; Sequence 9, Application US/08257341
; Patent No. 5525491
;  GENERAL INFORMATION:
;    APPLICANT:  HUSTON, JAMES S
;    APPLICANT:  OPPERMANN, HERMANN
;    APPLICANT:  TIMASHEFF, SERGE N
;    TITLE OF INVENTION:  SERINE RICH PEPTIDE LINKER
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CREATIVE BIOMOLECULES, INC./PATENT DEPT.
;      STREET:  35 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/257,341
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/842,149
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/662,226
;      FILING DATE:  27-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAMPBELL ESQ, PAULA A
;      REGISTRATION NUMBER:  32,503
;      REFERENCE/DOCKET NUMBER:  CRP-064CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/248-7000 (ATTY)
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  471 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-257-341-9

Alignment Scores:
Pred. No.:              23.1          Length:        471
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     1             Gaps:          0

US-10-677-865A-2 (1-1590) x US-08-257-341-9 (1-471)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||
Db        195 SerGlySerSerSerSerGlySerSer 203

RESULT 57
US-09-702-705-1812
; Sequence 1812, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C14
;  CURRENT APPLICATION NUMBER: US/09/702,705
;  CURRENT FILING DATE:  2000-10-30
;  NUMBER OF SEQ ID NOS:  1833
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1812
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-702-705-1812

Alignment Scores:
Pred. No.:              23.1          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-702-705-1812 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 58
US-09-736-457-1812
; Sequence 1812, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1812
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-736-457-1812

Alignment Scores:
Pred. No.:              23.1          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
```

```
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-736-457-1812 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 59
US-09-671-325-1812
; Sequence 1812, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1812
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-671-325-1812

Alignment Scores:
Pred. No.:              23.1          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-671-325-1812 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 60
US-09-538-092-1279
; Sequence 1279, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1279
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number Q06945
US-09-538-092-1279

Alignment Scores:
Pred. No.:              23.1          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-538-092-1279 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 61
US-10-017-754-1812
; Sequence 1812, Application US/10017754
; Patent No. 6858204
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1812
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-017-754-1812

Alignment Scores:
Pred. No.:              23.1          Length:        474
Score:                  9.00          Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            1.7%          Indels:        0
DB:                     2             Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-017-754-1812 (1-474)

Qy        454 AGCTCTCATTCGAGCAGCAGCAGCAGC 480
              ||||||||||||||||||||||||||||
Db        383 SerSerHisSerSerSerSerSerSer 391

RESULT 62
US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
;  APPLICANT: Ross, Jeffrey
;  TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
;  TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
;  FILE REFERENCE: 960296.95131
;  CURRENT APPLICATION NUMBER: US/09/261,855A
;  CURRENT FILING DATE: 1999-03-03
;  NUMBER OF SEQ ID NOS: 46
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 577
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.:                22.6          Length:          577
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       2             Gaps:            0

US-10-677-865A-2 (1-1590) x US-09-261-855-2 (1-577)

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAG 1521
              ||||||||||||||||||||||||||||
Db        553 IleArgAspIleLeuAlaGlnValLys 561

RESULT 63
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. 6794151
; GENERAL INFORMATION:
;  APPLICANT: Ross, Jeffrey
;  TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
;  TITLE OF INVENTION:  (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
;  FILE REFERENCE: 960296.95131
;  CURRENT APPLICATION NUMBER: US/09/873,637
;  CURRENT FILING DATE:  2001-06-04
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 577
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-873-637-2

Alignment Scores:
Pred. No.:                22.6          Length:          577
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       2             Gaps:            0

US-10-677-865A-2 (1-1590) x US-09-873-637-2 (1-577)

Qy       1495 ATCCGGGATATCCTAGCCCAAGTGAAG 1521
              ||||||||||||||||||||||||||||
Db        553 IleArgAspIleLeuAlaGlnValLys 561

RESULT 64
US-09-543-681A-4959
; Sequence 4959, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4959
;   LENGTH: 621
;   TYPE: PRT
;   ORGANISM: Proteus mirabilis
US-09-543-681A-4959

Alignment Scores:
Pred. No.:                22.4          Length:          621
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       2             Gaps:            0

US-10-677-865A-2 (1-1590) x US-09-543-681A-4959 (1-621)

Qy       1280 GCACCGGCTCCTTCAGCAGCAGCTCCA 1306
              ||||||||||||||||||||||||||||
Db         88 AlaProAlaProSerAlaAlaAlaPro 96

RESULT 65
US-08-356-786-16
; Sequence 16, Application US/08356786
; Patent No. 5877305
;  GENERAL INFORMATION:
;    APPLICANT:  Huston, James S.
;    APPLICANT:  Oppermann, Hermann
;    APPLICANT:  Houston, L. L.
;    APPLICANT:  Ring, David B.
;    TITLE OF INVENTION:  Biosynthetic Binding Protein for Cancer
;    TITLE OF INVENTION:  Marker
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Edmund R. Pitcher, Testa, Hurwitz, & Thibeault
;      STREET:  Exchange Place, 53 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/356,786
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/831,967
;      FILING DATE:  06-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;      REGISTRATION NUMBER:  27,829
;      REFERENCE/DOCKET NUMBER:  CRP-053
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  622 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-356-786-16

Alignment Scores:
Pred. No.:                22.4          Length:          622
Score:                    9.00          Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              1.7%          Indels:          0
DB:                       1             Gaps:            0

US-10-677-865A-2 (1-1590) x US-08-356-786-16 (1-622)

Qy        205 AGTGGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||||||
```

```
Db        125 SerGlySerSerSerSerGlySerSer 133

RESULT 66
US-08-986-837-6
; Sequence 6, Application US/08986837
; Patent No. 6221676
; GENERAL INFORMATION:
;  APPLICANT: Lam, Bing K.
;  APPLICANT: Penrose, John F.
;  APPLICANT: Frank, Austen K.
;  TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,
;  TITLE OF INVENTION: POLYPEPTIDES, AND USES THEREFOR
;  FILE REFERENCE: Docket No. 6221676 0092662-0019 (BWH269)
;  CURRENT APPLICATION NUMBER: US/08/986,837
;  CURRENT FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 08/246,991
;  EARLIER FILING DATE: 1994-05-20
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 24
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-986-837-6

Alignment Scores:
Pred. No.:              236          Length:         24
Score:                  8.00         Matches:        8
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            1.5%         Indels:         0
DB:                     2            Gaps:           0

US-10-677-865A-2 (1-1590) x US-08-986-837-6 (1-24)

Qy        137 CGCGTATCACCTCCCCTAACGACC 160
              ||||||||||||||||||||||||
Db         10 ArgValSerProProLeuThrThr 17

RESULT 67
US-09-809-592-6
; Sequence 6, Application US/09809592
; Patent No. 6576753
;   GENERAL INFORMATION:
;        APPLICANT: Austen M.D., K. Frank
;                   Lam Ph.D., Bing K.
;                   Penrose M.D., John F.
;        TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4
;                            SYNTHASE, POLYPEPTIDES, AND USES THEREFOR
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: CHOATE, HALL & STEWART
;             STREET: 53 State Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02109-2891
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/809,592
;             FILING DATE: 15-Mar-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jarrell Ph.D., Brenda H.
;             REGISTRATION NUMBER: 39,223
;             REFERENCE/DOCKET NUMBER: 0092662-0031
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 248-5000
;             TELEFAX: (617) 248-4000
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 24 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        FRAGMENT TYPE: internal
;        IMMEDIATE SOURCE:
;             CLONE: human LTC4 synthase loop residues 25 to 58
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-809-592-6

Alignment Scores:
Pred. No.:              236          Length:         24
Score:                  8.00         Matches:        8
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            1.5%         Indels:         0
DB:                     2            Gaps:           0

US-10-677-865A-2 (1-1590) x US-09-809-592-6 (1-24)

Qy        137 CGCGTATCACCTCCCCTAACGACC 160
              ||||||||||||||||||||||||
Db         10 ArgValSerProProLeuThrThr 17

RESULT 68
US-09-509-593-16
; Sequence 16, Application US/09509593
; Patent No. 6572860
;  GENERAL INFORMATION:
;    APPLICANT:  DANIEL H. ZIMMERMAN, ET AL.
;    TITLE OF INVENTION:  CONJUGATED POLYPEPTIDES FOR TREATING HERPES
;    TITLE OF INVENTION:  SIMPLEX VIRUS, COMPOSITIONS AND VACCINES CONTAINING SAME AND
;    TITLE OF INVENTION:  THEREOF FOR TREATMENT OR PREVENTION OF HERPES SIMPLEX VIRUS A
;    TITLE OF INVENTION:  DIAGNOSTIC TOOL AND IN DIAGNOSTIC KIT
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Law Office of Sherman and Shalloway
;      STREET:  413 N. Washington Street
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22313
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.25 inch, 1.4 mb storage
;      COMPUTER:  Dell Optiplex GX110; 733 Mhz Microprocessor
;      OPERATING SYSTEM:  Windows 98
;      SOFTWARE:  Word 2000
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/509,593
;      FILING DATE:  March 30, 2000
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Richard A. Steinberg
;      REGISTRATION NUMBER:  26,588
;      REFERENCE/DOCKET NUMBER:  CELL-104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 549-2282
;      TELEFAX:  (703) 836-0106
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acids
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;    HYPOTHETICAL:
;    ORIGINAL SOURCE:
;      ORGANISM:
;    FEATURE:
```

```
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
US-09-509-593-16

Alignment Scores:
Pred. No.:              233          Length:        27
Score:                  8.00         Matches:       8
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.5%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-509-593-16 (1-27)

Qy        253 GGTGGTGGCTCCAGTGGTAGCTCC 276
              ||||||||||||||||||||||||
Db         14 GlyGlyGlySerSerGlySerSer 21

RESULT 69
US-09-270-767-57471
; Sequence 57471, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57471
;   LENGTH: 34
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-57471

Alignment Scores:
Pred. No.:              228          Length:        34
Score:                  8.00         Matches:       8
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.5%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-270-767-57471 (1-34)

Qy        208 GGCTCCAGCAGTTCTGGCAGCTCC 231
              ||||||||||||||||||||||||
Db          9 GlySerSerSerSerGlySerSer 16

RESULT 70
US-09-461-325-306
; Sequence 306, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029P1
;  CURRENT APPLICATION NUMBER: US/09/461,325A
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: PCT/US99/13418
;  EARLIER FILING DATE: 1999-06-15
;  EARLIER APPLICATION NUMBER: 60/089,507
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,508
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,509
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,510
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/090,112
;  EARLIER FILING DATE: 1998-06-22
;  EARLIER APPLICATION NUMBER: 60/090,113
;  EARLIER FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
;   LENGTH: 39
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-461-325-306

Alignment Scores:
Pred. No.:              224          Length:        39
Score:                  8.00         Matches:       8
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.5%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x US-09-461-325-306 (1-39)

Qy       1302 GCTGCTGCTGAAGGAGCCGGTGCC 1279
              ||||||||||||||||||||||||
Db         16 AlaAlaAlaGluGlyAlaGlyAla 23

RESULT 71
US-10-012-542-306
; Sequence 306, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029P1
;  CURRENT APPLICATION NUMBER: US/10/012,542
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
;   LENGTH: 39
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-542-306

Alignment Scores:
Pred. No.:              224          Length:        39
Score:                  8.00         Matches:       8
Percent Similarity:     100.0%       Conservative:  0
Best Local Similarity:  100.0%       Mismatches:    0
Query Match:            1.5%         Indels:        0
DB:                     2            Gaps:          0

US-10-677-865A-2 (1-1590) x US-10-012-542-306 (1-39)

Qy       1302 GCTGCTGCTGAAGGAGCCGGTGCC 1279
              ||||||||||||||||||||||||
Db         16 AlaAlaAlaGluGlyAlaGlyAla 23

RESULT 72
US-10-115-123-306
; Sequence 306, Application US/10115123
```

; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 306
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-123-306

Alignment Scores:

Pred. No.:	224	Length:	39
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-10-115-123-306 (1-39)

```
Qy       1302 GCTGCTGCTGAAGGAGCCGGTGCC 1279
              ||||||||||||||||||||||||
Db         16 AlaAlaAlaGluGlyAlaGlyAla 23
```

RESULT 73
US-09-248-796A-27841
; Sequence 27841, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27841
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27841

Alignment Scores:

Pred. No.:	214	Length:	62
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-248-796A-27841 (1-62)

```
Qy       1198 AGCAGTTCTAGCATTTCCAGCAGC 1221
              ||||||||||||||||||||||||
Db          4 SerSerSerSerIleSerSerSer 11
```

RESULT 74
US-09-140-749-50
; Sequence 50, Application US/09140749
; Patent No. 6482795
; GENERAL INFORMATION:
; APPLICANT: Steck, Peter
; APPLICANT: Pershouse, Mark A.
; APPLICANT: Jasser, Samar
; APPLICANT: Yung, Alfred W.K.
; APPLICANT: Tavtigian, Sean V.
; TITLE OF INVENTION: A TUMOR SUPPRESSOR DESIGNATED TS10Q23.3
; FILE REFERENCE: 2318-205
; CURRENT APPLICATION NUMBER: US/09/140,749
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: U.S. 08/791,115
; EARLIER FILING DATE: 1997-01-30
; EARLIER APPLICATION NUMBER: U.S. 60/057,750
; EARLIER FILING DATE: 1997-08-26
; EARLIER APPLICATION NUMBER: U.S. 60/083,563
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-140-749-50

Alignment Scores:

Pred. No.:	210	Length:	75
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0
DB:	2	Gaps:	0

US-10-677-865A-2 (1-1590) x US-09-140-749-50 (1-75)

```
Qy       1305 GGAGCTGCTGCTGAAGGAGCCGGT 1282
              ||||||||||||||||||||||||
Db         28 GlyAlaAlaAlaGluGlyAlaGly 35
```

RESULT 75
US-09-198-452A-1151
; Sequence 1151, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1151
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1151

Alignment Scores:

Pred. No.:	206	Length:	90
Score:	8.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	1.5%	Indels:	0

DB: 2 Gaps: 0

US-10-677-865A-2 (1-1590) x US-09-198-452A-1151 (1-90)

```
Qy          58 CTGCTGCTGGCTGGTCTCCTCCTG 81
               ||||||||||||||||||||||||
Db          40 LeuLeuLeuAlaGlyLeuLeuLeu 47
```

Search completed: March 22, 2006, 02:07:33
Job time : 60.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:46:13 ; Search time 10.7636 Seconds
(without alignments)
143.025 Million cell updates/sec

Title: US-10-677-865A-3
Perfect score: 84
Sequence: 1 KSYGGYEVVGGSSDSG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	92.9	502	2	A48679	differentiated ker
2	48	57.1	260	2	C69130	shikimate 5-dehydr
3	46	54.8	289	2	A45369	insect-selective n
4	45	53.6	138	2	E83470	hypothetical prote
5	45	53.6	291	2	S16950	toxin TxP-I precur
6	45	53.6	413	2	S04655	transcription fact
7	44	52.4	295	2	A23750	30K actin-bundling
8	44	52.4	398	2	C81313	methionine adenosy
9	44	52.4	571	2	S58356	pepT protein - Sta
10	44	52.4	581	2	E83790	ABC transporter (A
11	43	51.2	208	2	H89776	hypothetical prote
12	43	51.2	358	2	E95842	probable lactose t
13	43	51.2	361	2	F96034	probable sugar upt
14	43	51.2	539	2	E87676	ATP synthase F1, b
15	43	51.2	539	2	AB3168	hypothetical prote
16	43	51.2	777	1	TVVPCP	large T antigen -
17	43	51.2	782	2	S22560	large T antigen -
18	43	51.2	785	1	TVVPT	large T antigen -
19	42	50.0	415	2	S43484	heterogeneous nucl
20	42	50.0	500	2	AD1047	probable amino aci
21	42	50.0	500	2	A86112	probable transport
22	42	50.0	514	2	S56384	hypothetical 56.3K
23	42	50.0	514	2	A91271	probable transport
24	42	50.0	548	2	A70980	probable accD5 pro
25	42	50.0	1041	2	E70760	probable ileS prot
26	42	50.0	2960	2	A45259	desmoyokin - human
27	41.5	49.4	668	2	T44118	penicillin-binding
28	41.5	49.4	668	2	JQ0773	penicillin-binding
29	41.5	49.4	668	2	JQ0774	penicillin-binding
30	41	48.8	227	2	T15772	hypothetical prote
31	41	48.8	258	1	QREBPT	histidine transpor
32	41	48.8	258	2	AE0800	histidine transpor
33	41	48.8	368	2	H96912	ABC-type probable
34	41	48.8	380	2	AE2321	S-adenosylmethioni
35	41	48.8	396	2	C95088	S-adenosylmethioni
36	41	48.8	396	2	G97955	methionine adenosy
37	41	48.8	399	2	F86862	methionine adenosy
38	41	48.8	409	2	S74736	methionine adenosy
39	41	48.8	413	2	S48756	transcription fact
40	41	48.8	496	2	T19776	hypothetical prote
41	41	48.8	526	1	KRBOVI	keratin, 54K type
42	41	48.8	561	2	T22917	probable ATP-depen
43	41	48.8	835	2	T47521	respiratory burst
44	41	48.8	1037	2	T18365	polysialyltransfer
45	41	48.8	1037	2	T18368	polysialyltransfer

ALIGNMENTS

RESULT 1
A48679
differentiated keratinocyte S protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A48679
R;Zhou, Y.; Chaplin, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 9470-9474, 1993
A;Title: Identification in the HLA class I region of a gene expressed late in keratinoc
A;Reference number: A48679; MUID:94022396; PMID:8415725
A;Accession: A48679
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <ZHO>
A;Cross-references: UNIPROT:O43509; UNIPARC:UPI0000177376; GB:L20815
A;Note: sequence shown follows the authors' translation at positions 47 and 410
C;Genetics:
A;Map position: 6p21.3
C;Superfamily: loricrin

Query Match 92.9%; Score 78; DB 2; Length 502;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KSYGGYEVVGGSSDS 15
              |||||||||||||||
Db        291 KSYGGYEVVGGSSDS 305
```

RESULT 2
C69130
shikimate 5-dehydrogenase - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C;Accession: C69130
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69130
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-260 <MTH>
A;Cross-references: UNIPARC:UPI0000165B85; GB:AE000811; GB:AE000666; NID:g2621287; PIDN
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH242
C;Superfamily: shikimate 5-dehydrogenase; shikimate dehydrogenase homology
F;39-240/Domain: shikimate dehydrogenase homology <SKD>

```
  Query Match             57.1%;  Score 48;  DB 2;  Length 260;
  Best Local Similarity   64.3%;  Pred. No. 4.6;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDS 15
              ||||||:: ||   |
Db        152 SYGGYELIQGSVKS 165


RESULT 3
A45369
insect-selective neurotoxin TxP-I homolog - straw itch mite
C;Species: Pyemotes tritici (straw itch mite)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A45369
R;Tomalski, M.D.; Hutchinson, K.; Todd, J.; Miller, L.K.
Toxicon 31, 319-326, 1993
A;Title: Identification and characterization of tox21A: a mite cDNA encoding a paralytic
A;Reference number: A45369; MUID:93227269; PMID:8470135
A;Accession: A45369
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-289 <TOM>
A;Cross-references: UNIPROT:Q07338; UNIPARC:UPI000013784C; GB:S58065; NID:g299121; PID:g
A;Note: sequence extracted from NCBI backbone (NCBIN:129517, NCBIP:129520)
C;Superfamily: straw itch mite toxin TxP-I

  Query Match             54.8%;  Score 46;  DB 2;  Length 289;
  Best Local Similarity   46.2%;  Pred. No. 10;
  Matches    6;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSS 13
              | |||::::|| :
Db         80 KVYGGFDIIGGKA 92


RESULT 4
E83470
hypothetical protein PA1395 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83470
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83470
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-138 <STO>
A;Cross-references: UNIPROT:Q9I3U9; UNIPARC:UPI00000C5358; GB:AE004569; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1395

  Query Match             53.6%;  Score 45;  DB 2;  Length 138;
  Best Local Similarity   69.2%;  Pred. No. 7;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSD 14
              ||  ||||| ||:
Db         69 SYSRYEVVGSSSE 81


RESULT 5
S16950
toxin TxP-I precursor - straw itch mite
C;Species: Pyemotes tritici (straw itch mite)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S16950; A33576
R;Tomalski, M.D.; Miller, L.K.
Nature 352, 82-85, 1991
A;Title: Insect paralysis by baculovirus-mediated expression of a mite neurotoxin gene.
A;Reference number: S16950; MUID:91287828; PMID:1840646
A;Accession: S16950
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-291 <TOM>
A;Cross-references: UNIPROT:P20798; UNIPARC:UPI000014ED17
R;Tomalski, M.D.; Kutney, R.; Bruce, W.A.; Brown, M.R.; Blum, M.S.; Travis, J.
Toxicon 27, 1151-1167, 1989
A;Title: Purification and characterization of insect toxins derived from the mite, Pyemc
A;Reference number: A33576; MUID:90050217; PMID:2815110
A;Accession: A33576
A;Status: preliminary
A;Molecule type: protein
A;Residues: 40-54,'X',56-60 <TO2>
A;Cross-references: UNIPARC:UPI00001766D0
C;Superfamily: straw itch mite toxin TxP-I
C;Keywords: neurotoxin

  Query Match             53.6%;  Score 45;  DB 2;  Length 291;
  Best Local Similarity   54.5%;  Pred. No. 15;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              | |||::::||
Db         87 KVYGGFDIIGG 97


RESULT 6
S04655
transcription factor GATA-1 - mouse
N;Alternate names: finger protein GF-1; protein EF-1; protein EF-gamma-a; protein Eryf1,
C;Species: Mus musculus (house mouse)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S04655; S55394
R;Tsai, S.F.; Martin, D.I.K.; Zon, L.I.; D'Andrea, A.D.; Wong, G.G.; Orkin, S.H.
Nature 339, 446-451, 1989
A;Title: Cloning of cDNA for the major DNA-binding protein of the erythroid lineage thro
A;Reference number: S04655; MUID:89262083; PMID:2725678
A;Accession: S04655
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-413 <TSA>
A;Cross-references: UNIPROT:P17679; UNIPARC:UPI00000236CF; GB:X15763; NID:g51063; PIDN:(
R;Todokoro, K.; Chiba, T.; Kuramochi, S.; Ikawa, Y.
submitted to the EMBL Data Library, February 1991
A;Description: Nucleotide sequences of the 5'-flanking region of the mouse GATA-1 gene i
A;Reference number: S55394
A;Accession: S55394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-73 <TOD>
A;Cross-references: UNIPARC:UPI000016CD78; EMBL:X57530; NID:g854279; PIDN:CAA40751.1; PI
C;Superfamily: transcription factor GATA-1; GATA-type zinc finger homology
C;Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc i
F;201-254/Domain: GATA-type zinc finger homology <GZF1>
F;204-228/Region: zinc finger GATA motif
F;255-308/Domain: GATA-type zinc finger homology <GZF2>
F;258-282/Region: zinc finger GATA motif

  Query Match             53.6%;  Score 45;  DB 2;  Length 413;
  Best Local Similarity   69.2%;  Pred. No. 22;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GGYEVVGGSSDSG 16
              ||: || ||| ||
Db        332 GGFMVVAGSSSSG 344


RESULT 7
A23750
30K actin-bundling protein - slime mold (Dictyostelium discoideum)
```

```
C;Species: Dictyostelium discoideum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Oct-2004
C;Accession: A23750
R;Fechheimer, M.; Murdock, D.; Carney, M.; Glover, C.V.C.
J. Biol. Chem. 266, 2883-2889, 1991
A;Title: Isolation and sequencing of cDNA clones encoding the Dictyostelium discoideum 3
A;Reference number: A23750; MUID:91131582; PMID:1993662
A;Accession: A23750
A;Molecule type: mRNA
A;Residues: 1-295 <FEC>
A;Cross-references: UNIPROT:P24005; UNIPARC:UPI0000125439; GB:M58022; NID:g167577; PIDN:
C;Superfamily: calcium-regulated actin bundling protein
C;Keywords: actin binding

  Query Match             52.4%;  Score 44;  DB 2;  Length 295;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSSDSG 16
              | |||    ||: |:|
Db        253 KKYGGAAYSGGAGDAG 268


RESULT 8
C81313
methionine adenosyltransferase (EC 2.5.1.6) Cj1096c [imported] - Campylobacter jejuni (s
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81313
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: C81313
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <PAR>
A;Cross-references: UNIPROT:Q9PNJ8; UNIPARC:UPI00000C1E14; GB:AL139077; GB:AL111168; NID
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: metK; Cj1096c
C;Superfamily: methionine adenosyltransferase
C;Keywords: S-adenosylmethionine; transferase

  Query Match             52.4%;  Score 44;  DB 2;  Length 398;
  Best Local Similarity   50.0%;  Pred. No. 30;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDS 15
              |:|||  :|| : |
Db        262 SFGGYSPIGGGAQS 275


RESULT 9
S58356
pepT protein - Staphylococcus epidermidis
C;Species: Staphylococcus epidermidis
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S66651; S58356
R;Meyer, C.; Bierbaum, G.; Heidrich, C.; Reis, M.; Sueling, J.; Iglesias-Wind, M.I.; Kem
Eur. J. Biochem. 232, 478-489, 1995
A;Title: Nucleotide sequence of the lantibiotic Pep5 biosynthetic gene cluster and funct
A;Reference number: S66651; MUID:96035882; PMID:7556197
A;Accession: S66651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-571 <ME2>
A;Cross-references: UNIPROT:Q54121; UNIPARC:UPI00000B7E58; EMBL:Z49865; NID:g945014; PID
C;Genetics:
A;Gene: pepT
C;Superfamily: Escherichia coli ABC transporter mdlA; ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop
F;348-543/Domain: ATP-binding cassette homology <ABC>
F;365-372/Region: nucleotide-binding motif A (P-loop)

  Query Match             52.4%;  Score 44;  DB 2;  Length 571;
  Best Local Similarity   53.3%;  Pred. No. 44;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDSG 16
              ||||| :  |: ||
Db        259 SYGGYRIAVGAITSG 273


RESULT 10
E83790
ABC transporter (ATP-binding protein) BH1125 [imported] - Bacillus halodurans (strain C
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: E83790
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: E83790
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-581 <STO>
A;Cross-references: UNIPROT:Q9KDT6; UNIPARC:UPI00000C3ACF; GB:AP001511; GB:BA000004; NI
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1125
C;Superfamily: Escherichia coli ABC transporter mdlA; ATP-binding cassette homology

  Query Match             52.4%;  Score 44;  DB 2;  Length 581;
  Best Local Similarity   53.3%;  Pred. No. 45;
  Matches    8;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SYGGYEVVGGSSDSG 16
              |:||| |: ||  :|
Db        259 SFGGYLVITGSMQTG 273


RESULT 11
H89776
hypothetical protein capH [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H89776
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: H89776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-208 <KUR>
A;Cross-references: UNIPROT:P72386; UNIPARC:UPI00000D472C; GB:BA000018; PID:g13700072;
A;Experimental source: strain N315
C;Genetics:
A;Gene: capH
C;Superfamily: Agrobacterium chloramphenicol acetyltransferase

  Query Match             51.2%;  Score 43;  DB 2;  Length 208;
  Best Local Similarity   72.7%;  Pred. No. 22;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              |: | ||||||
Db        154 KNVGAYEVVGG 164


RESULT 12
```

```
E95842
probable lactose transport ATP-binding protein lacK1 [imported] - Sinorhizobium meliloti
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: E95842
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: E95842
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-358 <KUR>
A;Cross-references: UNIPROT:Q92XF6; UNIPARC:UPI00000CB39A; GB:AL591985; PIDN:CAC48405.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: lacK1; SMb20002
A;Genome: plasmid
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology

  Query Match             51.2%;  Score 43;  DB 2;  Length 358;
  Best Local Similarity   72.7%;  Pred. No. 39;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              |||||  ||: |
Db         11 KSYGGLEVIKG 21


RESULT 13
F96034
probable sugar uptake ABC transporter ATP-binding protein SMb20630 [imported] - Sinorhiz
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: F96034
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F96034
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <KUR>
A;Cross-references: UNIPROT:Q92TH6; UNIPARC:UPI00000CB8D1; GB:AL591985; PIDN:CAC49942.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20630
A;Genome: plasmid
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology

  Query Match             51.2%;  Score 43;  DB 2;  Length 361;
  Best Local Similarity   72.7%;  Pred. No. 39;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              |||||  ||: |
Db         17 KSYGGLEVIHG 27


RESULT 14
E87676
ATP synthase F1, beta subunit [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 31-Dec-2004
C;Accession: E87676
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87676
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-539 <STO>
A;Cross-references: UNIPROT:Q9A2V9; UNIPARC:UPI00000C7AA0; GB:AE005673; NID:g13425165; P
C;Genetics:
A;Gene: CC3447
C;Superfamily: H(+)-transporting ATP synthase; H+-transporting ATP synthase alpha chain

  Query Match             51.2%;  Score 43;  DB 2;  Length 539;
  Best Local Similarity   63.6%;  Pred. No. 59;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGG 11
              |:|||| |: |
Db        231 KAYGGYSVLAG 241


RESULT 15
AB3168
hypothetical protein Atu5068 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 31-Dec-2004
C;Accession: AB3168
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB3168
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-539 <KUR>
A;Cross-references: UNIPROT:Q8UKN2; UNIPARC:UPI00000D14BC; GB:AE008687; PIDN:AAL45760.1,
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu5068
A;Genome: plasmid

  Query Match             51.2%;  Score 43;  DB 2;  Length 539;
  Best Local Similarity   61.5%;  Pred. No. 59;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 KSYGGYEVVGGSS 13
              |||||:| : | |
Db        283 KSYGGFEALKGIS 295


Search completed: March 21, 2006, 21:49:27
Job time : 12.7636 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:43 ; Search time 32.1818 Seconds
(without alignments)
136.530 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	486	7	ABJ39147	Abj39147 Human S p
2	56	100.0	486	8	ADQ21504	Adq21504 Human sof
3	56	100.0	529	2	AAW80498	Aaw80498 A protein
4	56	100.0	529	2	AAW85055	Aaw85055 Human pro
5	56	100.0	529	3	AAY44989	Aay44989 Human epi
6	56	100.0	529	4	AAG64152	Aag64152 Polypepti
7	56	100.0	529	4	AAU27550	Aau27550 Corneodes
8	56	100.0	529	7	ABJ39148	Abj39148 Human cor
9	56	100.0	529	8	ADN04226	Adn04226 Antipsori
10	56	100.0	529	8	ADP12525	Adp12525 Protein e
11	46	82.1	249	5	AAU81187	Aau81187 Proteorho
12	46	82.1	249	8	ADQ82668	Adq82668 Mutant ma
13	46	82.1	249	8	ADQ82670	Adq82670 Mutant ma
14	46	82.1	249	8	ADQ82508	Adq82508 Wild type
15	46	82.1	249	8	ADQ82530	Adq82530 Wild type
16	46	82.1	249	8	ADQ82534	Adq82534 Wild type
17	46	82.1	249	8	ADQ82672	Adq82672 Mutant ma
18	46	82.1	249	8	ADQ82610	Adq82610 Wild type
19	46	82.1	250	5	AAU81198	Aau81198 Proteorho
20	46	82.1	250	5	AAU81186	Aau81186 Proteorho
21	46	82.1	250	5	AAU81196	Aau81196 Proteorho
22	46	82.1	250	5	AAU81197	Aau81197 Proteorho
23	46	82.1	250	5	AAU81202	Aau81202 Proteorho
24	46	82.1	250	5	AAU81204	Aau81204 Proteorho
25	46	82.1	250	5	AAU81191	Aau81191 Proteorho
26	46	82.1	250	5	AAU81188	Aau81188 Proteorho
27	46	82.1	250	5	AAU81206	Aau81206 Proteorho
28	46	82.1	250	5	AAU81201	Aau81201 Proteorho
29	46	82.1	250	5	AAU81203	Aau81203 Proteorho
30	46	82.1	250	5	AAU81199	Aau81199 Proteorho
31	46	82.1	250	5	AAU81205	Aau81205 Proteorho
32	46	82.1	250	5	AAU81207	Aau81207 Proteorho
33	46	82.1	250	5	AAU81189	Aau81189 Proteorho
34	46	82.1	250	5	AAU81200	Aau81200 Proteorho
35	46	82.1	250	8	ADQ82604	Adq82604 Wild type
36	46	82.1	250	8	ADQ82574	Adq82574 Wild type
37	46	82.1	250	8	ADQ82600	Adq82600 Wild type
38	46	82.1	250	8	ADQ82572	Adq82572 Wild type
39	46	82.1	250	8	ADQ82584	Adq82584 Wild type
40	46	82.1	250	8	ADQ82602	Adq82602 Wild type
41	46	82.1	250	8	ADQ82622	Adq82622 Wild type
42	46	82.1	250	8	ADQ82564	Adq82564 Wild type
43	46	82.1	250	8	ADQ82566	Adq82566 Wild type
44	46	82.1	250	8	ADQ82606	Adq82606 Wild type
45	46	82.1	250	8	ADQ82608	Adq82608 Wild type

ALIGNMENTS

RESULT 1
ABJ39147
ID ABJ39147 standard; protein; 486 AA.
XX
AC ABJ39147;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human S protein encoded by the skin corneodesmosin (CDSN) gene (S-gene).
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6;
KW S protein; skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
DR N-PSDB; ABT43848.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2b; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to

```
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human S protein (corneodesmosin protein only
CC   expressed in the skin), the target of the invention
XX
SQ   Sequence 486 AA;

  Query Match            100.0%;  Score 56;  DB 7;  Length 486;
  Best Local Similarity  100.0%;  Pred. No. 0.075;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        301 KIYPVGYFTK 310


RESULT 2
ADQ21504
ID   ADQ21504 standard; protein; 486 AA.
XX
AC   ADQ21504;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 4324.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 4324; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 486 AA;

  Query Match            100.0%;  Score 56;  DB 8;  Length 486;
  Best Local Similarity  100.0%;  Pred. No. 0.075;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        301 KIYPVGYFTK 310


RESULT 3
AAW80498
ID   AAW80498 standard; protein; 529 AA.
XX
AC   AAW80498;
XX
DT   10-FEB-1999  (first entry)
XX
DE   A protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761362-A1.
XX
PD   02-OCT-1998.
XX
PF   28-MAR-1997;   97FR-00003899.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586093/50.
DR   N-PSDB; AAV63614.
XX
PT   New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT   or therapeutic treatment of trophic disorders or thinning of the
PT   epidermis.
XX
PS   Claim 1; Page 29; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

  Query Match            100.0%;  Score 56;  DB 2;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 4
AAW85055
ID   AAW85055 standard; protein; 529 AA.
XX
AC   AAW85055;
XX
DT   05-FEB-1999  (first entry)
```

```
XX
DE   Human protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761363-A1.
XX
PD   02-OCT-1998.
XX
PF   11-SEP-1997;   97FR-00011317.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586094/50.
DR   N-PSDB; AAV63541.
XX
PT   New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT   cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT   epidermis etc., also compositions containing protease that degrades
PT   corneodesmosin.
XX
PS   Claim 1; Page 30; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 5
AAY44989
ID   AAY44989 standard; protein; 529 AA.
XX
AC   AAY44989;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human epidermal protein-6.
XX
KW   Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW   dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW   arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .32
FT                   /label= Signal_peptide
FT   Modified-site   3
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   37
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   50
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   57
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   80
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   172
FT                   /note= "N-glycosylated"
FT   Modified-site   182
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   287
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   395
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   440
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   N-PSDB; AAZ50582.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 1; Fig 4; 82pp; English.
XX
CC   The present sequence is human epidermal protein-6 (HEPI) expressed in
CC   breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC   cDNA library. HEPI-6 shares 99% identity with human S protein.
CC   Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC   for protein expression. The HEPI proteins are useful for the treatment of
CC   epithelial disorders, including dyshidrotic eczema and scabies, cell
CC   proliferative disorders including actinic keratosis and arteriosclerosis,
CC   and autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 3;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
```

```
Db        317 KIYPVGYFTK 326


RESULT 6
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 143
FT                   /label= Ser, Asn
FT   Misc-difference 151
FT                   /label= Ser, Cys
FT   Misc-difference 152
FT                   /label= Ser, Arg
FT   Misc-difference 200
FT                   /label= Ser, Phe
FT   Misc-difference 282
FT                   /label= Pro, Thr
FT   Misc-difference 401
FT                   /label= Ser, Gly
FT   Misc-difference 408
FT                   /label= Ala, Ser
FT   Misc-difference 409
FT                   /label= Gly, Val
FT   Misc-difference 410
FT                   /label= Ser, Leu
FT   Misc-difference 453
FT                   /label= Ser, Asn
FT   Misc-difference 527
FT                   /label= Asp, Asn
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 7
AAU27550
ID   AAU27550 standard; protein; 529 AA.
XX
AC   AAU27550;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
DR   N-PSDB; AAS43749.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Claim 8; Fig 2; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. The present sequence represents corneodesmosin amino acid
CC   sequence, as described in the invention
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326
```

```
RESULT 8
ABJ39148
ID   ABJ39148 standard; protein; 529 AA.
XX
AC   ABJ39148;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human corneodesmosin (CDSN) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 28
FT                   /note= "Encoded by CCAG"
FT   misc_difference 408
FT                   /note= "Encoded by TCC"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   N-PSDB; ABT43849, ABT43850.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 7;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326
```

```
RESULT 9
ADN04226
ID   ADN04226 standard; protein; 529 AA.
XX
AC   ADN04226;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #308.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04225.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 8;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326
```

```
RESULT 10
ADP12525
ID   ADP12525 standard; protein; 529 AA.
XX
AC   ADP12525;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #135.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
```

```
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J,  Fry K,  Woodward R,  Ly N,  Prentice J,  Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   protein that is encoded by the mRNA of the invention.
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 56;  DB 8;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 11
AAU81187
ID   AAU81187 standard; protein; 249 AA.
XX
AC   AAU81187;
XX
DT   07-AUG-2003  (revised)
DT   09-APR-2002  (first entry)
XX
DE   Proteorhodopsin from clone EBAC31A8 #2.
XX
KW   Proteorhodopsin; light-driven energy generator; targeted drug delivery;
KW   biocatalytic reactor; fuel cell; nano-machine; molecular switching;
KW   data storage; membrane potential; halophilic arachaea.
XX
OS   Eubacteria.
XX
PN   WO200183701-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014394.
XX
PR   03-MAY-2000; 2000US-0201602P.
XX
PA   (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI   Delong EF,  Beja O;
XX
DR   WPI; 2002-114151/15.
DR   N-PSDB; ABK24126.
XX
PT   New proteorhodopsin gene for use in a light-driven energy generator,
PT   comprises an isolated DNA sequence encoding a proteorhodopsin protein.
XX
PS   Disclosure; Fig 5; 460pp; English.
XX
CC   The invention describes a proteorhodopsin gene, comprising an isolated
CC   DNA sequence for encoding a proteorhodopsin protein. The gene is useful
CC   in light-driven energy generators. A proteorhodopsin system is useful in
CC   many industrial and technological applications, for use in targeted drug
CC   delivery, has primary and secondary energy generators for biocatalytic
CC   reactors, fuel cells and nano-machines, as well as uses in molecular
CC   switching or data storage devices. Proteorhodopsin is useful for a
CC   process to enhance yield or increase the potential of recombinant protein
CC   production or convert the light-induced membrane potential into cellular
CC   signals. The system is not restricted to operate in halophilic arachaea
CC   and could therefore be functionally expressed in Escherichia coli and
CC   other bacteria. It provides for a fast and cheap production method that
CC   allows for mass production of functionally active proteorhodopsin. This
CC   is the amino acid sequence of a proteorhodopsin protein variant,
CC   described in the method of the invention. (Updated on 07-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 249 AA;

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


RESULT 12
ADQ82668
ID   ADQ82668 standard; protein; 249 AA.
XX
AC   ADQ82668;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Mutant marine eubacteria proteorhodopsin protein #1.
XX
KW   proteorhodopsin; improved optical characteristic;
KW   integral membrane protein; marine eubacteria; light-driven proton pump;
KW   photochromic all-trans-retinal protein; optical storage device;
KW   information storage; holographic storage; associative storage;
KW   information processing; optical bistability; light switching;
KW   optical filtering; signal conditioning; neural networks;
KW   spatial light modulator; phase conjugation; pattern recognition;
KW   interferometry; ATP generation; sea water desalination; mutein; mutant.
XX
OS   unidentified marine eubacterium.
OS   Synthetic.
XX
PN   WO2004063326-A2.
XX
PD   29-JUL-2004.
XX
PF   26-NOV-2003; 2003WO-US038194.
XX
PR   26-NOV-2002; 2002US-0429518P.
XX
PA   (GEMV ) GENENCOR INT INC.
PA   (DOWO ) DOW CORNING CORP.
XX
PI   Jensen RB,  Kelemen BR;
XX
DR   WPI; 2004-544077/52.
DR   N-PSDB; ADQ82669.
XX
```

PT New proteorhodopsin mutant having improved optical characteristics,
PT useful as a photochromic all-trans-retinal protein for optical storage
PT devices, or in devices for ATP generation in reactors and desalination of
PT sea water.
XX
PS Claim 14; SEQ ID NO 163; 316pp; English.
XX
CC The invention comprises the amino acid and coding sequences of mutant
CC proteorhodopsin proteins which have improved optical characteristics.
CC Proteorhodopsins are integral membrane proteins from marine eubacteria
CC which function as light-driven proton pumps. The mutant proteorhodopsin
CC proteins of the invention are useful for as photochromic all-trans-
CC retinal proteins for optical storage devices, in devices for information
CC storage such as 2-D storage, 3-D storage, holographic storage, or
CC associative storage. The mutant proteorhodopsin proteins are useful in
CC devices for information processing such as optical bistability/light
CC switching, optical filtering, signal conditioning, neural networks,
CC spatial light modulators, phase conjugation, pattern recognition, and
CC interferometry. The mutant proteorhodopsin proteins are also useful in
CC devices for ATP generation in reactors, desalination of sea water, and/or
CC conversion of sunlight into electricity. The present amino acid sequence
CC represents a mutant marine eubacteria proteorhodopsin protein of the
CC invention.
XX
SQ Sequence 249 AA;

Query Match 82.1%; Score 46; DB 8; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IYPVGYFT 9
||||||||
Db 199 IYPVGYFT 206

RESULT 13
ADQ82670
ID ADQ82670 standard; protein; 249 AA.
XX
AC ADQ82670;
XX
DT 21-OCT-2004 (first entry)
XX
DE Mutant marine eubacteria proteorhodopsin protein #2.
XX
KW proteorhodopsin; improved optical characteristic;
KW integral membrane protein; marine eubacteria; light-driven proton pump;
KW photochromic all-trans-retinal protein; optical storage device;
KW information storage; holographic storage; associative storage;
KW information processing; optical bistability; light switching;
KW optical filtering; signal conditioning; neural networks;
KW spatial light modulator; phase conjugation; pattern recognition;
KW interferometry; ATP generation; sea water desalination; mutein; mutant.
XX
OS unidentified marine eubacterium.
OS Synthetic.
XX
PN WO2004063326-A2.
XX
PD 29-JUL-2004.
XX
PF 26-NOV-2003; 2003WO-US038194.
XX
PR 26-NOV-2002; 2002US-0429518P.
XX
PA (GEMV) GENENCOR INT INC.
PA (DOWO) DOW CORNING CORP.
XX
PI Jensen RB, Kelemen BR;
XX
DR WPI; 2004-544077/52.
DR N-PSDB; ADQ82671.
XX

PT New proteorhodopsin mutant having improved optical characteristics,
PT useful as a photochromic all-trans-retinal protein for optical storage
PT devices, or in devices for ATP generation in reactors and desalination of
PT sea water.
XX
PS Claim 14; SEQ ID NO 165; 316pp; English.
XX
CC The invention comprises the amino acid and coding sequences of mutant
CC proteorhodopsin proteins which have improved optical characteristics.
CC Proteorhodopsins are integral membrane proteins from marine eubacteria
CC which function as light-driven proton pumps. The mutant proteorhodopsin
CC proteins of the invention are useful for as photochromic all-trans-
CC retinal proteins for optical storage devices, in devices for information
CC storage such as 2-D storage, 3-D storage, holographic storage, or
CC associative storage. The mutant proteorhodopsin proteins are useful in
CC devices for information processing such as optical bistability/light
CC switching, optical filtering, signal conditioning, neural networks,
CC spatial light modulators, phase conjugation, pattern recognition, and
CC interferometry. The mutant proteorhodopsin proteins are also useful in
CC devices for ATP generation in reactors, desalination of sea water, and/or
CC conversion of sunlight into electricity. The present amino acid sequence
CC represents a mutant marine eubacteria proteorhodopsin protein of the
CC invention.
XX
SQ Sequence 249 AA;

Query Match 82.1%; Score 46; DB 8; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IYPVGYFT 9
||||||||
Db 199 IYPVGYFT 206

RESULT 14
ADQ82508
ID ADQ82508 standard; protein; 249 AA.
XX
AC ADQ82508;
XX
DT 21-OCT-2004 (first entry)
XX
DE Wild type marine eubacteria proteorhodopsin protein #2.
XX
KW proteorhodopsin; improved optical characteristic;
KW integral membrane protein; marine eubacteria; light-driven proton pump;
KW photochromic all-trans-retinal protein; optical storage device;
KW information storage; holographic storage; associative storage;
KW information processing; optical bistability; light switching;
KW optical filtering; signal conditioning; neural networks;
KW spatial light modulator; phase conjugation; pattern recognition;
KW interferometry; ATP generation; sea water desalination.
XX
OS unidentified marine eubacterium.
XX
PN WO2004063326-A2.
XX
PD 29-JUL-2004.
XX
PF 26-NOV-2003; 2003WO-US038194.
XX
PR 26-NOV-2002; 2002US-0429518P.
XX
PA (GEMV) GENENCOR INT INC.
PA (DOWO) DOW CORNING CORP.
XX
PI Jensen RB, Kelemen BR;
XX
DR WPI; 2004-544077/52.
DR N-PSDB; ADQ82509.
XX
PT New proteorhodopsin mutant having improved optical characteristics,

PT useful as a photochromic all-trans-retinal protein for optical storage
PT devices, or in devices for ATP generation in reactors and desalination of
PT sea water.
XX
PS Claim 4; SEQ ID NO 3; 316pp; English.
XX
CC The invention comprises the amino acid and coding sequences of mutant
CC proteorhodopsin proteins which have improved optical characteristics.
CC Proteorhodopsins are integral membrane proteins from marine eubacteria
CC which function as light-driven proton pumps. The mutant proteorhodopsin
CC proteins of the invention are useful for as photochromic all-trans-
CC retinal proteins for optical storage devices, in devices for information
CC storage such as 2-D storage, 3-D storage, holographic storage, or
CC associative storage. The mutant proteorhodopsin proteins are useful in
CC devices for information processing such as optical bistability/light
CC switching, optical filtering, signal conditioning, neural networks,
CC spatial light modulators, phase conjugation, pattern recognition, and
CC interferometry. The mutant proteorhodopsin proteins are also useful in
CC devices for ATP generation in reactors, desalination of sea water, and/or
CC conversion of sunlight into electricity. The present amino acid sequence
CC represents a wild type marine eubacteria proteorhodopsin protein of the
CC invention.
XX
SQ Sequence 249 AA;

Query Match 82.1%; Score 46; DB 8; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IYPVGYFT 9
||||||||
Db 199 IYPVGYFT 206

RESULT 15
ADQ82530
ID ADQ82530 standard; protein; 249 AA.
XX
AC ADQ82530;
XX
DT 21-OCT-2004 (first entry)
XX
DE Wild type marine eubacteria proteorhodopsin protein #14.
XX
KW proteorhodopsin; improved optical characteristic;
KW integral membrane protein; marine eubacteria; light-driven proton pump;
KW photochromic all-trans-retinal protein; optical storage device;
KW information storage; holographic storage; associative storage;
KW information processing; optical bistability; light switching;
KW optical filtering; signal conditioning; neural networks;
KW spatial light modulator; phase conjugation; pattern recognition;
KW interferometry; ATP generation; sea water desalination.
XX
OS unidentified marine eubacterium.
XX
PN WO2004063326-A2.
XX
PD 29-JUL-2004.
XX
PF 26-NOV-2003; 2003WO-US038194.
XX
PR 26-NOV-2002; 2002US-0429518P.
XX
PA (GEMV) GENENCOR INT INC.
PA (DOWO) DOW CORNING CORP.
XX
PI Jensen RB, Kelemen BR;
XX
DR WPI; 2004-544077/52.
DR N-PSDB; ADQ82531.
XX
PT New proteorhodopsin mutant having improved optical characteristics,
PT useful as a photochromic all-trans-retinal protein for optical storage
PT devices, or in devices for ATP generation in reactors and desalination of
PT sea water.
XX
PS Claim 4; SEQ ID NO 25; 316pp; English.
XX
CC The invention comprises the amino acid and coding sequences of mutant
CC proteorhodopsin proteins which have improved optical characteristics.
CC Proteorhodopsins are integral membrane proteins from marine eubacteria
CC which function as light-driven proton pumps. The mutant proteorhodopsin
CC proteins of the invention are useful for as photochromic all-trans-
CC retinal proteins for optical storage devices, in devices for information
CC storage such as 2-D storage, 3-D storage, holographic storage, or
CC associative storage. The mutant proteorhodopsin proteins are useful in
CC devices for information processing such as optical bistability/light
CC switching, optical filtering, signal conditioning, neural networks,
CC spatial light modulators, phase conjugation, pattern recognition, and
CC interferometry. The mutant proteorhodopsin proteins are also useful in
CC devices for ATP generation in reactors, desalination of sea water, and/or
CC conversion of sunlight into electricity. The present amino acid sequence
CC represents a wild type marine eubacteria proteorhodopsin protein of the
CC invention.
XX
SQ Sequence 249 AA;

Query Match 82.1%; Score 46; DB 8; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IYPVGYFT 9
||||||||
Db 201 IYPVGYFT 208

Search completed: March 21, 2006, 21:45:55
Job time : 32.1818 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:43 ; Search time 26.1818 Seconds
(without alignments)
159.588 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	486	4	US-10-316-175-12	Sequence 12, Appl
2	56	100.0	486	5	US-10-723-860-4324	Sequence 4324, Ap
3	56	100.0	486	5	US-10-883-805-16	Sequence 16, Appl
4	56	100.0	529	4	US-10-164-230-61	Sequence 61, Appl
5	56	100.0	529	4	US-10-204-884-3	Sequence 3, Appli
6	56	100.0	529	4	US-10-316-175-15	Sequence 15, Appl
7	56	100.0	529	5	US-10-883-805-6	Sequence 6, Appli
8	56	100.0	529	5	US-10-677-865-1	Sequence 1, Appli
9	46	82.1	249	3	US-09-847-513A-7	Sequence 7, Appli
10	46	82.1	249	5	US-10-724-264A-3	Sequence 3, Appli
11	46	82.1	249	5	US-10-724-264A-25	Sequence 25, Appl
12	46	82.1	249	5	US-10-724-264A-29	Sequence 29, Appl
13	46	82.1	249	5	US-10-724-264A-105	Sequence 105, App
14	46	82.1	249	5	US-10-724-264A-163	Sequence 163, App
15	46	82.1	249	5	US-10-724-264A-165	Sequence 165, App
16	46	82.1	249	5	US-10-724-264A-167	Sequence 167, App
17	46	82.1	250	3	US-09-847-513A-5	Sequence 5, Appli
18	46	82.1	250	3	US-09-847-513A-9	Sequence 9, Appli
19	46	82.1	250	3	US-09-847-513A-11	Sequence 11, Appl
20	46	82.1	250	3	US-09-847-513A-15	Sequence 15, Appl
21	46	82.1	250	3	US-09-847-513A-25	Sequence 25, Appl
22	46	82.1	250	3	US-09-847-513A-27	Sequence 27, Appl
23	46	82.1	250	3	US-09-847-513A-29	Sequence 29, Appl
24	46	82.1	250	3	US-09-847-513A-31	Sequence 31, Appl
25	46	82.1	250	3	US-09-847-513A-33	Sequence 33, Appl
26	46	82.1	250	3	US-09-847-513A-35	Sequence 35, Appl
27	46	82.1	250	3	US-09-847-513A-37	Sequence 37, Appl
28	46	82.1	250	3	US-09-847-513A-39	Sequence 39, Appl
29	46	82.1	250	3	US-09-847-513A-41	Sequence 41, Appl
30	46	82.1	250	3	US-09-847-513A-43	Sequence 43, Appl
31	46	82.1	250	3	US-09-847-513A-45	Sequence 45, Appl
32	46	82.1	250	3	US-09-847-513A-47	Sequence 47, Appl
33	46	82.1	250	5	US-10-724-264A-59	Sequence 59, Appl
34	46	82.1	250	5	US-10-724-264A-61	Sequence 61, Appl
35	46	82.1	250	5	US-10-724-264A-63	Sequence 63, Appl
36	46	82.1	250	5	US-10-724-264A-65	Sequence 65, Appl
37	46	82.1	250	5	US-10-724-264A-67	Sequence 67, Appl
38	46	82.1	250	5	US-10-724-264A-69	Sequence 69, Appl
39	46	82.1	250	5	US-10-724-264A-71	Sequence 71, Appl
40	46	82.1	250	5	US-10-724-264A-75	Sequence 75, Appl
41	46	82.1	250	5	US-10-724-264A-77	Sequence 77, Appl
42	46	82.1	250	5	US-10-724-264A-79	Sequence 79, Appl
43	46	82.1	250	5	US-10-724-264A-81	Sequence 81, Appl
44	46	82.1	250	5	US-10-724-264A-83	Sequence 83, Appl
45	46	82.1	250	5	US-10-724-264A-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-316-175-12
; Sequence 12, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-12

Query Match 100.0%; Score 56; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIYPVGYFTK 10
||||||||||
Db 301 KIYPVGYFTK 310

RESULT 2
US-10-723-860-4324
; Sequence 4324, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4324
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-4324

  Query Match             100.0%;  Score 56;  DB 5;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        301 KIYPVGYFTK 310


RESULT 3
US-10-883-805-16
; Sequence 16, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 16
;   LENGTH: 486
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No: g414810
US-10-883-805-16

  Query Match             100.0%;  Score 56;  DB 5;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 0.077;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        301 KIYPVGYFTK 310


RESULT 4
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             100.0%;  Score 56;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 5
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
;  APPLICANT: Oxagen Limited
;  APPLICANT:  Olaveson, Mark
;  APPLICANT:  Lench, Nick
;  APPLICANT:  Allen, Maxine
```

```
;  APPLICANT:  Tazi-Ahnini, Rachind
;  TITLE OF INVENTION: Test and model for inflammatory disease
;  FILE REFERENCE: P30000WO-PS
;  CURRENT APPLICATION NUMBER: US/10/204,884
;  CURRENT FILING DATE:  2003-01-29
;  PRIOR APPLICATION NUMBER: GB 0004312.5
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 189
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (143)..(143)
;   OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (401)..(401)
;   OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (408)..(408)
;   OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (409)..(409)
;   OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (410)..(410)
;   OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (527)..(527)
;   OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

  Query Match             100.0%;  Score 56;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 6
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
;  APPLICANT: WELLA AG
;  APPLICANT:  Pras, Elon
;  APPLICANT:  Levy-Nissenbaum, Etgar
;  APPLICANT:  Frydman, Moshe
;  APPLICANT:  Thederahn, Ted
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING HYPOTRICHOSIS SIMPLEX
;  FILE REFERENCE: 532862000100
;  CURRENT APPLICATION NUMBER: US/10/316,175
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: 60/338,188
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-15

  Query Match             100.0%;  Score 56;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 7
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

  Query Match             100.0%;  Score 56;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 8
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
;  APPLICANT: SERRE, Guy Bruno Rene
;  APPLICANT:  SIMON, Michel
```

```
;  APPLICANT:  WEBER-VIVAT, Marina
;  TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
;  TITLE OF INVENTION:  AND USE THEREOF
;  FILE REFERENCE: 016800-336
;  CURRENT APPLICATION NUMBER: US/10/677,865
;  CURRENT FILING DATE:  2003-10-03
;  PRIOR APPLICATION NUMBER: US/09/381,656
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: FR 97/03899
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-10-677-865-1

  Query Match             100.0%;  Score 56;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.084;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 9
US-09-847-513A-7
; Sequence 7, Application US/09847513A
; Publication No. US20030104375A1
; GENERAL INFORMATION:
;  APPLICANT: MBARI
;  APPLICANT:  DeLong, Edward
;  APPLICANT:  Beja, Oded
;  TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
;  FILE REFERENCE: MBA-101
;  CURRENT APPLICATION NUMBER: US/09/847,513A
;  CURRENT FILING DATE:  2001-05-01
;  PRIOR APPLICATION NUMBER: 60/201,602
;  PRIOR FILING DATE: 2000-05-03
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Naturally occurring gamma proteobacterium
US-09-847-513A-7

  Query Match             82.1%;  Score 46;  DB 3;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


RESULT 10
US-10-724-264A-3
; Sequence 3, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-3

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


RESULT 11
US-10-724-264A-25
; Sequence 25, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-25

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        201 IYPVGYFT 208


RESULT 12
US-10-724-264A-29
; Sequence 29, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-29

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
```

```
              ||||||||
Db        201 IYPVGYFT 208


RESULT 13
US-10-724-264A-105
; Sequence 105, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 105
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-105

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        200 IYPVGYFT 207


RESULT 14
US-10-724-264A-163
; Sequence 163, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 163
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-163

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


RESULT 15
US-10-724-264A-165
; Sequence 165, Application US/10724264A
; Publication No. US20050095605A1
; GENERAL INFORMATION:
;  APPLICANT: JENSEN, Rasmus B.
;  APPLICANT:  KELEMEN, Bradley
;  TITLE OF INVENTION: PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS
;  FILE REFERENCE: 02716.0005.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/724,264A
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,518
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 165
;   LENGTH: 249
;   TYPE: PRT
;   ORGANISM: Marine eubacteria
US-10-724-264A-165

  Query Match             82.1%;  Score 46;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFT 9
              ||||||||
Db        199 IYPVGYFT 206


Search completed: March 21, 2006, 21:52:53
Job time : 27.1818 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:50:33 ; Search time 4.36364 Seconds
(without alignments)
65.594 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37	66.1	188	6	US-10-821-234-941	Sequence 941, App
2	37	66.1	404	7	US-11-087-099-11850	Sequence 11850, A
3	34	60.7	16	7	US-11-054-515-2173	Sequence 2173, Ap
4	34	60.7	19	7	US-11-054-515-2936	Sequence 2936, Ap
5	34	60.7	251	7	US-11-054-515-856	Sequence 856, App
6	34	60.7	253	7	US-11-054-515-1520	Sequence 1520, Ap
7	34	60.7	256	7	US-11-054-515-1463	Sequence 1463, Ap
8	34	60.7	759	7	US-11-096-568A-29706	Sequence 29706, A
9	34	60.7	764	7	US-11-096-568A-29705	Sequence 29705, A
10	34	60.7	767	7	US-11-096-568A-29704	Sequence 29704, A
11	34	60.7	1515	7	US-11-124-367A-402	Sequence 402, App
12	34	60.7	1516	7	US-11-124-367A-401	Sequence 401, App
13	34	60.7	1518	7	US-11-124-367A-406	Sequence 406, App
14	34	60.7	1532	7	US-11-124-367A-403	Sequence 403, App
15	34	60.7	1532	7	US-11-124-367A-404	Sequence 404, App
16	34	60.7	1532	7	US-11-124-367A-405	Sequence 405, App
17	34	60.7	1532	7	US-11-124-367A-407	Sequence 407, App
18	34	60.7	1535	7	US-11-124-367A-408	Sequence 408, App
19	33.5	59.8	702	7	US-11-043-788-371	Sequence 371, App
20	33.5	59.8	727	7	US-11-043-788-367	Sequence 367, App
21	33.5	59.8	727	7	US-11-043-788-368	Sequence 368, App
22	33.5	59.8	729	7	US-11-043-788-373	Sequence 373, App
23	33.5	59.8	736	7	US-11-043-788-372	Sequence 372, App
24	33.5	59.8	774	7	US-11-043-788-370	Sequence 370, App
25	33	58.9	159	6	US-10-467-657-2908	Sequence 2908, Ap
26	33	58.9	1177	7	US-11-226-943-10	Sequence 10, Appl
27	33	58.9	1177	7	US-11-226-943-12	Sequence 12, Appl
28	33	58.9	1177	7	US-11-226-943-14	Sequence 14, Appl
29	33	58.9	1177	7	US-11-226-943-26	Sequence 26, Appl
30	33	58.9	1177	7	US-11-226-943-28	Sequence 28, Appl
31	33	58.9	1177	7	US-11-226-943-34	Sequence 34, Appl
32	33	58.9	2764	6	US-10-995-561-691	Sequence 691, App
33	33	58.9	2813	6	US-10-995-561-688	Sequence 688, App
34	33	58.9	2919	6	US-10-821-234-1133	Sequence 1133, Ap
35	32	57.1	117	7	US-11-049-536-642	Sequence 642, App
36	32	57.1	117	7	US-11-049-536-658	Sequence 658, App
37	32	57.1	121	7	US-11-049-536-202	Sequence 202, App
38	32	57.1	127	7	US-11-049-536-294	Sequence 294, App
39	32	57.1	134	6	US-10-993-543-40	Sequence 40, Appl
40	32	57.1	287	7	US-11-092-168-5	Sequence 5, Appli
41	32	57.1	308	6	US-10-330-773-650	Sequence 650, App
42	32	57.1	314	7	US-11-096-568A-11685	Sequence 11685, A
43	32	57.1	329	7	US-11-096-568A-11684	Sequence 11684, A
44	32	57.1	397	7	US-11-096-568A-11683	Sequence 11683, A
45	32	57.1	497	7	US-11-072-512-2347	Sequence 2347, Ap

ALIGNMENTS

RESULT 1
US-10-821-234-941
; Sequence 941, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 941
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-941

Query Match 66.1%; Score 37; DB 6; Length 188;
Best Local Similarity 66.7%; Pred. No. 7;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 IYPVGYFTK 10
              :|||||  |
Db         47 LYPVGYLDK 55
```

RESULT 2
US-11-087-099-11850
; Sequence 11850, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11850
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bradyrhizobium japonicum USDA 110
US-11-087-099-11850

```
  Query Match             66.1%;  Score 37;  DB 7;  Length 404;
  Best Local Similarity   77.8%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              | | |||||
Db        204 IPPAGYFTK 212


RESULT 3
US-11-054-515-2173
; Sequence 2173, Application US/11054515
; Publication No. US20050255532A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523P3
;  CURRENT APPLICATION NUMBER: US/11/054,515
;  CURRENT FILING DATE:  2005-02-10
;  PRIOR APPLICATION NUMBER: 60/543,296
;  PRIOR FILING DATE: 2004-02-11
;  PRIOR APPLICATION NUMBER: 60/580,347
;  PRIOR FILING DATE: 2004-06-18
;  PRIOR APPLICATION NUMBER: 10/293,418
;  PRIOR FILING DATE: 2002-11-14
;  PRIOR APPLICATION NUMBER: 60/331,469
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/340,817
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 09/880,748
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2173
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-054-515-2173

  Query Match             60.7%;  Score 34;  DB 7;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 2;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YPVGYF 8
              ||:|||
Db          9 YPMGYF 14


RESULT 4
US-11-054-515-2936
; Sequence 2936, Application US/11054515
; Publication No. US20050255532A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523P3
;  CURRENT APPLICATION NUMBER: US/11/054,515
;  CURRENT FILING DATE:  2005-02-10
;  PRIOR APPLICATION NUMBER: 60/543,296
;  PRIOR FILING DATE: 2004-02-11
;  PRIOR APPLICATION NUMBER: 60/580,347
;  PRIOR FILING DATE: 2004-06-18
;  PRIOR APPLICATION NUMBER: 10/293,418
;  PRIOR FILING DATE: 2002-11-14
;  PRIOR APPLICATION NUMBER: 60/331,469
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/340,817
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 09/880,748
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2936
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-054-515-2936

  Query Match             60.7%;  Score 34;  DB 7;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 2.4;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YPVGYF 8
              ||:|||
Db         12 YPLGYF 17


RESULT 5
US-11-054-515-856
; Sequence 856, Application US/11054515
; Publication No. US20050255532A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523P3
;  CURRENT APPLICATION NUMBER: US/11/054,515
;  CURRENT FILING DATE:  2005-02-10
;  PRIOR APPLICATION NUMBER: 60/543,296
;  PRIOR FILING DATE: 2004-02-11
;  PRIOR APPLICATION NUMBER: 60/580,347
;  PRIOR FILING DATE: 2004-06-18
;  PRIOR APPLICATION NUMBER: 10/293,418
;  PRIOR FILING DATE: 2002-11-14
;  PRIOR APPLICATION NUMBER: 60/331,469
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/340,817
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 09/880,748
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 856
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-054-515-856

  Query Match             60.7%;  Score 34;  DB 7;  Length 251;
  Best Local Similarity   83.3%;  Pred. No. 35;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          3 YPVGYF 8
              ||:|||
Db        107 YPMGYF 112


RESULT 6
US-11-054-515-1520
; Sequence 1520, Application US/11054515
; Publication No. US20050255532A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523P3
;  CURRENT APPLICATION NUMBER: US/11/054,515
;  CURRENT FILING DATE:  2005-02-10
;  PRIOR APPLICATION NUMBER: 60/543,296
;  PRIOR FILING DATE: 2004-02-11
;  PRIOR APPLICATION NUMBER: 60/580,347
;  PRIOR FILING DATE: 2004-06-18
;  PRIOR APPLICATION NUMBER: 10/293,418
;  PRIOR FILING DATE: 2002-11-14
;  PRIOR APPLICATION NUMBER: 60/331,469
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/340,817
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 09/880,748
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1520
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-054-515-1520

  Query Match             60.7%;  Score 34;  DB 7;  Length 253;
  Best Local Similarity   83.3%;  Pred. No. 35;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YPVGYF 8
              ||:|||
Db        107 YPMGYF 112


RESULT 7
US-11-054-515-1463
; Sequence 1463, Application US/11054515
; Publication No. US20050255532A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Antibodies that Immunospecifically Bind BLyS
;  FILE REFERENCE: PF523P3
;  CURRENT APPLICATION NUMBER: US/11/054,515
;  CURRENT FILING DATE:  2005-02-10
;  PRIOR APPLICATION NUMBER: 60/543,296
;  PRIOR FILING DATE: 2004-02-11
;  PRIOR APPLICATION NUMBER: 60/580,347
;  PRIOR FILING DATE: 2004-06-18
;  PRIOR APPLICATION NUMBER: 10/293,418
;  PRIOR FILING DATE: 2002-11-14
;  PRIOR APPLICATION NUMBER: 60/331,469
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: 60/340,817
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: 09/880,748
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,379
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/240,816
;  PRIOR FILING DATE: 2000-10-17
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1463
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (49)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (53)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (54)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-054-515-1463

  Query Match             60.7%;  Score 34;  DB 7;  Length 256;
  Best Local Similarity   83.3%;  Pred. No. 35;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 YPVGYF 8
              ||:|||
Db        112 YPLGYF 117


RESULT 8
US-11-096-568A-29706
; Sequence 29706, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29706
;   LENGTH: 759
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(759)
;   OTHER INFORMATION: Ceres Seq. ID no. 4926017
US-11-096-568A-29706

  Query Match             60.7%;  Score 34;  DB 7;  Length 759;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYF 8
              :||:||:
Db         46 VYPLGYY 52


RESULT 9
US-11-096-568A-29705
; Sequence 29705, Application US/11096568A
; Publication No. US20060048240A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29705
;   LENGTH: 764
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(764)
;   OTHER INFORMATION: Ceres Seq. ID no. 4926016
US-11-096-568A-29705

  Query Match             60.7%;  Score 34;  DB 7;  Length 764;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYF 8
              :||:||:
Db         51 VYPLGYY 57


RESULT 10
US-11-096-568A-29704
; Sequence 29704, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29704
;   LENGTH: 767
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(767)
;   OTHER INFORMATION: Ceres Seq. ID no. 4926015
US-11-096-568A-29704

  Query Match             60.7%;  Score 34;  DB 7;  Length 767;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGYF 8
              :||:||:
Db         54 VYPLGYY 60


RESULT 11
US-11-124-367A-402
; Sequence 402, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 402
;   LENGTH: 1515
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-124-367A-402

  Query Match             60.7%;  Score 34;  DB 7;  Length 1515;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              ::|:|||  :
Db       1435 LWPIGYFLR 1443


RESULT 12
US-11-124-367A-401
; Sequence 401, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 401
;   LENGTH: 1516
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-124-367A-401

  Query Match             60.7%;  Score 34;  DB 7;  Length 1516;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              ::|:|||  :
Db       1436 LWPIGYFLR 1444


RESULT 13
US-11-124-367A-406
; Sequence 406, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
```

```
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 406
;    LENGTH: 1518
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-124-367A-406

  Query Match             60.7%;  Score 34;  DB 7;  Length 1518;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              ::|:||| :
Db       1438 LWPIGYFLR 1446


RESULT 14
US-11-124-367A-403
; Sequence 403, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 403
;    LENGTH: 1532
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-124-367A-403

  Query Match             60.7%;  Score 34;  DB 7;  Length 1532;
  Best Local Similarity   44.4%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              ::|:||| :
Db       1452 LWPIGYFLR 1460


RESULT 15
US-11-124-367A-404
; Sequence 404, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 404
;    LENGTH: 1532
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-124-367A-404

  Query Match             60.7%;  Score 34;  DB 7;  Length 1532;
  Best Local Similarity   44.4%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              ::|:||| :
Db       1452 LWPIGYFLR 1460


Search completed: March 21, 2006, 21:53:23
Job time : 5.36364 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:04 ; Search time 9.27273 Seconds
(without alignments)
89.160 Million cell updates/sec

Title: US-10-677-865A-4
Perfect score: 56
Sequence: 1 KIYPVGYFTK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	100.0	529	2	US-09-381-656-1	Sequence 1, Appli
2	56	100.0	529	2	US-10-164-230-61	Sequence 61, Appl
3	39	69.6	487	2	US-10-418-036-16	Sequence 16, Appl
4	37	66.1	214	2	US-09-205-258-1020	Sequence 1020, Ap
5	37	66.1	214	2	US-10-004-860-1020	Sequence 1020, Ap
6	37	66.1	220	2	US-09-489-039A-9187	Sequence 9187, Ap
7	37	66.1	248	2	US-09-949-016-7997	Sequence 7997, Ap
8	36	64.3	76	2	US-09-270-767-35717	Sequence 35717, A
9	36	64.3	76	2	US-09-270-767-50934	Sequence 50934, A
10	36	64.3	223	2	US-09-538-092-1038	Sequence 1038, Ap
11	36	64.3	772	2	US-09-543-681A-6467	Sequence 6467, Ap
12	35	62.5	177	2	US-09-270-767-33655	Sequence 33655, A
13	35	62.5	177	2	US-09-270-767-48872	Sequence 48872, A
14	35	62.5	195	2	US-09-198-452A-844	Sequence 844, App
15	35	62.5	280	2	US-09-107-532A-4849	Sequence 4849, Ap
16	35	62.5	294	2	US-09-438-185A-796	Sequence 796, App
17	34	60.7	117	2	US-09-543-681A-5626	Sequence 5626, Ap
18	34	60.7	152	2	US-09-270-767-47052	Sequence 47052, A
19	34	60.7	201	2	US-09-252-991A-27995	Sequence 27995, A
20	34	60.7	208	2	US-09-489-039A-11072	Sequence 11072, A
21	34	60.7	269	2	US-09-543-681A-6475	Sequence 6475, Ap
22	34	60.7	286	2	US-09-134-000C-4900	Sequence 4900, Ap
23	34	60.7	324	1	US-08-597-236-10	Sequence 10, Appl
24	34	60.7	324	1	US-08-746-682A-10	Sequence 10, Appl
25	34	60.7	397	1	US-08-774-104A-2	Sequence 2, Appli
26	34	60.7	397	2	US-09-511-532-2	Sequence 2, Appli
27	34	60.7	412	2	US-09-574-141A-60	Sequence 60, Appl
28	34	60.7	412	2	US-09-568-189A-60	Sequence 60, Appl
29	33.5	59.8	183	2	US-09-450-209-6	Sequence 6, Appli
30	33.5	59.8	727	2	US-08-989-385-4	Sequence 4, Appli
31	33.5	59.8	727	2	US-09-593-826-4	Sequence 4, Appli
32	33.5	59.8	736	2	US-09-949-016-8038	Sequence 8038, Ap
33	33	58.9	115	2	US-09-134-001C-2914	Sequence 2914, Ap
34	33	58.9	140	2	US-09-461-325-241	Sequence 241, App
35	33	58.9	140	2	US-09-461-325-424	Sequence 424, App
36	33	58.9	140	2	US-10-012-542-241	Sequence 241, App
37	33	58.9	140	2	US-10-012-542-424	Sequence 424, App
38	33	58.9	140	2	US-10-115-123-241	Sequence 241, App
39	33	58.9	140	2	US-10-115-123-424	Sequence 424, App
40	33	58.9	231	2	US-09-270-767-42303	Sequence 42303, A
41	33	58.9	298	2	US-09-134-001C-3482	Sequence 3482, Ap
42	33	58.9	301	2	US-09-134-001C-2915	Sequence 2915, Ap
43	33	58.9	309	2	US-10-029-180-98	Sequence 98, Appl
44	33	58.9	311	2	US-10-029-180-12	Sequence 12, Appl
45	33	58.9	330	2	US-09-543-681A-8337	Sequence 8337, Ap

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Query Match 100.0%; Score 56; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326
```

RESULT 2
US-10-164-230-61
; Sequence 61, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             100.0%;  Score 56;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 0.019;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ||||||||||
Db        317 KIYPVGYFTK 326


RESULT 3
US-10-418-036-16
; Sequence 16, Application US/10418036
; Patent No. 6893833
; GENERAL INFORMATION:
;  APPLICANT: Gronberg, Alvar
;  APPLICANT:  Wikstrom, Per
;  TITLE OF INVENTION: NEW USE
;  FILE REFERENCE: 13425-110001
;  CURRENT APPLICATION NUMBER: US/10/418,036
;  CURRENT FILING DATE:  2003-04-17
;  PRIOR APPLICATION NUMBER: SE 0201152-6
;  PRIOR FILING DATE: 2002-04-17
;  PRIOR APPLICATION NUMBER: US 60/410,626
;  PRIOR FILING DATE: 2002-09-13
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-418-036-16

  Query Match             69.6%;  Score 39;  DB 2;  Length 487;
  Best Local Similarity   77.8%;  Pred. No. 29;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFT 9
              |||| | ||
Db        319 KIYPTGLFT 327


RESULT 4
US-09-205-258-1020
; Sequence 1020, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/09/205,258
;  CURRENT FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: PCT/US98/11422
;  EARLIER FILING DATE: 1998-06-04
;  EARLIER APPLICATION NUMBER: 60/048,885
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,375
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,881
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,880
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,896
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,020
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,876
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,895
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,884
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,894
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,971
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,882
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,899
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,893
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,900
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,901
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,892
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,915
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,019
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
```

```
;  EARLIER APPLICATION NUMBER: 60/048,972
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,916
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,373
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,875
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/049,374
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,917
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,949
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,883
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,897
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,898
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,962
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,963
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,877
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/048,878
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/070,923
;  EARLIER FILING DATE: 1997-12-18
;  EARLIER APPLICATION NUMBER: 60/092,921
;  EARLIER FILING DATE: 1998-07-15
;  EARLIER APPLICATION NUMBER: 60/094,657
;  EARLIER FILING DATE: 1998-07-30
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1020
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-205-258-1020

  Query Match             66.1%;  Score 37;  DB 2;  Length 214;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              :|||||  |
Db         73 LYPVGYLDK 81


RESULT 5
US-10-004-860-1020
; Sequence 1020, Application US/10004860
; Patent No. 6914047
; GENERAL INFORMATION:
;  APPLICANT: Young et al.
;  TITLE OF INVENTION: 207 Human Secreted Proteins
;  FILE REFERENCE: PZ007P1
;  CURRENT APPLICATION NUMBER: US/10/004,860
;  CURRENT FILING DATE:  2001-12-07
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 1227
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1020
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-004-860-1020

  Query Match             66.1%;  Score 37;  DB 2;  Length 214;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              :|||||  |
Db         73 LYPVGYLDK 81


RESULT 6
US-09-489-039A-9187
; Sequence 9187, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 9187
;   LENGTH: 220
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9187

  Query Match             66.1%;  Score 37;  DB 2;  Length 220;
  Best Local Similarity   85.7%;  Pred. No. 29;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KIYPVGY 7
              :||||||
Db        214 RIYPVGY 220


RESULT 7
US-09-949-016-7997
; Sequence 7997, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7997
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7997

  Query Match             66.1%;  Score 37;  DB 2;  Length 248;
  Best Local Similarity   60.0%;  Pred. No. 33;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              ::|||| | |
Db         72 EVYPVGIFQK 81


RESULT 8
US-09-270-767-35717
```

```
; Sequence 35717, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35717
;   LENGTH: 76
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-35717

  Query Match             64.3%;  Score 36;  DB 2;  Length 76;
  Best Local Similarity   77.8%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              || |||| |
Db         50 IYAVGYFWK 58


RESULT 9
US-09-270-767-50934
; Sequence 50934, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50934
;   LENGTH: 76
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-50934

  Query Match             64.3%;  Score 36;  DB 2;  Length 76;
  Best Local Similarity   77.8%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 IYPVGYFTK 10
              || |||| |
Db         50 IYAVGYFWK 58


RESULT 10
US-09-538-092-1038
; Sequence 1038, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1038
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P30041
US-09-538-092-1038

  Query Match             64.3%;  Score 36;  DB 2;  Length 223;
  Best Local Similarity   60.0%;  Pred. No. 46;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              |::| | |||
Db        199 KLFPKGVFTK 208


RESULT 11
US-09-543-681A-6467
; Sequence 6467, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6467
;   LENGTH: 772
;   TYPE: PRT
;   ORGANISM: Proteus mirabilis
US-09-543-681A-6467

  Query Match             64.3%;  Score 36;  DB 2;  Length 772;
  Best Local Similarity   58.3%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          1 KIYPV--GYFTK 10
              |:||   |||:|
Db        757 KVYPTRKGYFSK 768


RESULT 12
US-09-270-767-33655
; Sequence 33655, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33655
;   LENGTH: 177
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-33655

  Query Match             62.5%;  Score 35;  DB 2;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGY 7
              ||||||
Db        138 IYPVGY 143


RESULT 13
US-09-270-767-48872
```

```
; Sequence 48872, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48872
;   LENGTH: 177
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-48872

  Query Match             62.5%;  Score 35;  DB 2;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IYPVGY 7
              ||||||
Db        138 IYPVGY 143


RESULT 14
US-09-198-452A-844
; Sequence 844, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;  APPLICANT: Griffais, R.
;  TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;  TITLE OF INVENTION:  thereof and uses thereof, in particular for the diagnosis, preve
;  TITLE OF INVENTION:  and treatment of infection
;  FILE REFERENCE: 9710-003-999
;  CURRENT APPLICATION NUMBER: US/09/198,452A
;  CURRENT FILING DATE:  1998-11-24
;  NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 844
;   LENGTH: 195
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-198-452A-844

  Query Match             62.5%;  Score 35;  DB 2;  Length 195;
  Best Local Similarity   75.0%;  Pred. No. 61;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 YPVGYFTK 10
              : ||||||
Db         75 FSVGYFTK 82


RESULT 15
US-09-107-532A-4849
; Sequence 4849, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 4849:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 280 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...280
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4849:
US-09-107-532A-4849

  Query Match             62.5%;  Score 35;  DB 2;  Length 280;
  Best Local Similarity   60.0%;  Pred. No. 92;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 KIYPVGYFTK 10
              || |:||  |
Db        132 KILPIGYIRK 141


Search completed: March 21, 2006, 21:50:23
Job time : 10.2727 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:46:13 ; Search time 19.5091 Seconds
(without alignments)
143.025 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	37.6	300	2	C70586	probable bex - Myc
2	52	34.9	154	2	D72725	hypothetical prote
3	51.5	34.6	837	2	S64785	hypothetical prote
4	51	34.2	265	2	JC4970	slow muscle tropon
5	50.5	33.9	211	2	S50686	hypothetical prote
6	50	33.6	1441	2	B86807	hypothetical prote
7	49	32.9	283	2	B87546	acetoin dehydrogen
8	49	32.9	530	2	B37471	capsid protein - N
9	49	32.9	891	2	T36423	probable large, mu
10	48.5	32.6	337	2	D64049	adhesin homolog HI
11	48.5	32.6	524	2	C69297	acetyl-CoA decarbo
12	48	32.2	288	2	D69117	nitrogenase iron-m
13	48	32.2	309	2	A83702	inosine-uridine nu
14	48	32.2	352	2	T13993	hypothetical prote
15	48	32.2	473	2	AC2251	hypothetical prote
16	47.5	31.9	285	2	S66076	transcription repr
17	47.5	31.9	423	2	H83603	hypothetical prote
18	47.5	31.9	563	2	T02720	beta-D-glucosidase
19	47.5	31.9	570	2	B96776	hypothetical prote
20	47	31.5	171	2	JE0153	mitochondrial inne
21	47	31.5	240	2	AD2006	glucose-6-P-dehydr
22	47	31.5	257	2	T10586	small nuclear ribo
23	47	31.5	297	2	S70552	calnexin homolog C
24	47	31.5	342	2	C75172	enolase related (e
25	47	31.5	435	2	T16721	hypothetical prote
26	46.5	31.2	229	2	H98224	hypothetical prote
27	46.5	31.2	229	2	AG3061	conserved hypothet
28	46.5	31.2	511	1	S60287	glucose-6-phosphat
29	46.5	31.2	667	2	S74254	homeotic protein s
30	46.5	31.2	1804	2	T34518	nestin - golden ha
31	46	30.9	167	2	A81973	hypothetical prote
32	46	30.9	226	2	H84132	beta-phosphoglucom
33	46	30.9	574	1	S76132	protein kinase pkn
34	46	30.9	673	2	S46520	luciferin-binding
35	46	30.9	2019	2	A33996	sodium channel pro
36	46	30.9	4957	2	T03455	ALR protein - huma
37	46	30.9	5262	2	T03454	ALR protein - huma
38	45.5	30.5	250	2	H82079	protein disulfide-
39	45.5	30.5	309	2	JC4906	calponin - human
40	45.5	30.5	359	2	T35179	vanillate O-demeth
41	45.5	30.5	391	2	S75649	renin-binding prot
42	45.5	30.5	566	2	A48860	beta-glucosidase,
43	45.5	30.5	948	2	S75991	hypothetical prote
44	45.5	30.5	1191	2	A53491	bumetanide-sensiti
45	45	30.2	205	2	A41315	transcription repr

ALIGNMENTS

RESULT 1
C70586
probable bex - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Dec-2004
C;Accession: C70586
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70586
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-300 <COL>
A;Cross-references: UNIPROT:O05834; UNIPARC:UPI000012A0F5; GB:Z95208; GB:AL123456; NID:
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: bex
C;Superfamily: GTP-binding protein era homolog, bacteria type; translation elongation f
C;Keywords: GTP binding; nucleotide binding; P-loop
F;7-128/Domain: translation elongation factor Tu homology <ETU>
F;13-20/Region: nucleotide-binding motif A (P-loop)

Query Match 37.6%; Score 56; DB 2; Length 300;
Best Local Similarity 41.4%; Pred. No. 2.5;
Matches 12; Conservative 2; Mismatches 9; Indels 6; Gaps 1;

```
Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        168 DVLAAALPAGP------AYYPDGELTDEP 190
```

RESULT 2
D72725
hypothetical protein APE0341 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: D72725
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: D72725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <KAW>
A;Cross-references: UNIPROT:Q9YFA0; UNIPARC:UPI000005DADD; DDBJ:AP000059; NID:g5103911;
A;Experimental source: strain K1

```
C;Genetics:
A;Gene: APE0341
C;Superfamily: Aeropyrum pernix hypothetical protein APE0341

  Query Match             34.9%;  Score 52;  DB 2;  Length 154;
  Best Local Similarity   36.7%;  Pred. No. 4.4;
  Matches   11;  Conservative    7;  Mismatches    8;  Indels    4;  Gaps    2;

Qy          3 LAQVKPLGPQLADPEVFLPQGELL---DSP 29
              ||  :||||     | :::| | ::   |:|
Db          7 LAAARPLGPSRV-PPIYIPVGSMIRHYDNP 35


RESULT 3
S64785
hypothetical protein YLL034c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L0919
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C;Accession: S64785
R;Duesterhoeft, A.; Floeth, M.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64775
A;Accession: S64785
A;Molecule type: DNA
A;Residues: 1-837 <DUE>
A;Cross-references: UNIPROT:Q07844; UNIPARC:UPI0000053235; EMBL:Z73139; NID:g1360221; PI
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0003957
A;Map position: 12L
C;Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind
C;Keywords: ATP; nucleotide binding; P-loop
F;219-431/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT1>
F;246-253/Region: nucleotide-binding motif A (P-loop)
F;547-758/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAT2>
F;574-581/Region: nucleotide-binding motif A (P-loop)

  Query Match             34.6%;  Score 51.5;  DB 2;  Length 837;
  Best Local Similarity   38.9%;  Pred. No. 35;
  Matches   14;  Conservative    6;  Mismatches    9;  Indels    7;  Gaps    2;

Qy          1 DILAQVKPL-GPQLADPEVFL------PQGELLDSP 29
              |::||:  | |  :  ||:||      |:| ||  |
Db        212 DVVAQLMELIGLPILHPEIFLSTGVEPPRGVLLHGP 247


RESULT 4
JC4970
slow muscle troponin T - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 09-Jul-2004
C;Accession: JC4970
R;Yonemura, I.; Watanabe, T.; Kirinoki, M.; Miyazaki, J.; Hirabayashi, T.
Biochem. Biophys. Res. Commun. 226, 200-205, 1996
A;Title: Cloning of chicken slow muscle troponin T and its sequence comparison with that
A;Reference number: JC4970; MUID:96400238; PMID:8806614
A;Accession: JC4970
A;Molecule type: mRNA
A;Residues: 1-265 <YON>
A;Cross-references: UNIPROT:Q98916; UNIPARC:UPI00000FCA02; DDBJ:D85105; NID:g1620885; PI
C;Genetics:
A;Gene: TnT
C;Superfamily: troponin T

  Query Match             34.2%;  Score 51;  DB 2;  Length 265;
  Best Local Similarity   47.6%;  Pred. No. 11;
  Matches   10;  Conservative    6;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          7 KPLGPQLADPEVFLPQGELLD 27
              :|: |||| |::  |:|| :|
Db         48 RPVVPQLAPPKI--PEGERVD 66


RESULT 5
S50686
hypothetical protein YER183c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: S50686
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9163 and 9132.
A;Reference number: S50431
A;Accession: S50686
A;Molecule type: DNA
A;Residues: 1-211 <DIE>
A;Cross-references: UNIPROT:P40099; UNIPARC:UPI000012AC87; EMBL:U18922; NID:g603405; PI
C;Genetics:
A;Gene: SGD:FAU1; MIPS:YER183c
A;Cross-references: SGD:S0000985
A;Map position: 5R

  Query Match             33.9%;  Score 50.5;  DB 2;  Length 211;
  Best Local Similarity   37.0%;  Pred. No. 10;
  Matches   10;  Conservative    8;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          2 ILAQVKPLGP-QLADPEVFLPQGELLD 27
              ::  :|| || || :||  : : ::||
Db        107 MVRDLKPQGPYQLKEPEPHIEESDILD 133


RESULT 6
B86807
hypothetical protein yoiC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86807
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrl
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86807
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1441 <STO>
A;Cross-references: UNIPROT:Q9CFL1; UNIPARC:UPI00000D4470; GB:AE005176; PID:g12724450;
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yoiC

  Query Match             33.6%;  Score 50;  DB 2;  Length 1441;
  Best Local Similarity   37.0%;  Pred. No. 1e+02;
  Matches   10;  Conservative    8;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          5 QVKPLGPQLADPE--VFLPQGELLDSP 29
              :||| || :::|:  : || |  : :|
Db       1233 KVKPNGPSISNPDGSITLPGGGTVTTP 1259


RESULT 7
B87546
acetoin dehydrogenase E2 component, probable [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87546
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <STO>
```

```
A;Cross-references: UNIPROT:Q9A5Q2; UNIPARC:UPI00000C76F6; GB:AE005673; NID:g13423930; P
C;Genetics:
A;Gene: CC2395
C;Superfamily: tropinesterase

  Query Match             32.9%;  Score 49;  DB 2;  Length 283;
  Best Local Similarity   35.1%;  Pred. No. 23;
  Matches   13;  Conservative    4;  Mismatches   12;  Indels    8;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLP--------QGELLDSP 29
              | :| :    | :  ||||||         |||:  |
Db        111 DRVAGLMLASPAMMGPEVFLPFRLMSLPLLGELMSKP 147


RESULT 8
B37471
capsid protein - Norwalk virus
C;Species: Norwalk virus
C;Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: B37471
R;Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
Virology 195, 51-61, 1993
A;Title: Sequence and genomic organization of Norwalk virus.
A;Reference number: A37471; MUID:93303939; PMID:8391187
A;Accession: B37471
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: genomic RNA
A;Residues: 1-530 <JIA>
A;Cross-references: UNIPROT:Q83884; UNIPARC:UPI00000ED272; GB:M87661; NID:g1061311; PIDN
A;Note: sequence extracted from NCBI backbone (NCBIP:134157)
C;Superfamily: human calicivirus capsid protein

  Query Match             32.9%;  Score 49;  DB 2;  Length 530;
  Best Local Similarity   44.0%;  Pred. No. 47;
  Matches   11;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          5 QVKPLGPQLADPEVFLPQGELLDSP 29
              || |: | : :  |  ||||   ||
Db         48 QVNPIDPWIINNFVQAPQGEFTISP 72


RESULT 9
T36423
probable large, multifunctional secreted protein - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36423
R;Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21606
A;Accession: T36423
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-891 <SAU>
A;Cross-references: UNIPROT:Q8CK54; UNIPARC:UPI000017ADF0; EMBL:AL109974; PIDN:CAB53335.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCF34.24

  Query Match             32.9%;  Score 49;  DB 2;  Length 891;
  Best Local Similarity   31.7%;  Pred. No. 85;
  Matches   13;  Conservative    7;  Mismatches    7;  Indels   14;  Gaps    2;

Qy          3 LAQVK----------PLGP----QLADPEVFLPQGELLDSP 29
              | |:|          | ||      : :| ::||| |: :||
Db        421 LVQIKQDRFFNHYTEPAGPFEDSPVTEPVLWLPQNEIGNSP 461


RESULT 10
D64049
adhesin homolog HI0119 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: D64049
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: D64049
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-337 <TIGR>
A;Cross-references: UNIPROT:P44526; UNIPARC:UPI000013C439; GB:U32698; GB:L42023; NID:g3
C;Superfamily: hypothetical protein HI0119

  Query Match             32.6%;  Score 48.5;  DB 2;  Length 337;
  Best Local Similarity   43.3%;  Pred. No. 34;
  Matches   13;  Conservative    4;  Mismatches    6;  Indels    7;  Gaps    2;

Qy          1 DILAQVKPLG---PQLAD----PEVFLPQG 23
              |:|| |||||     :||     :| :| | |
Db         24 DVLASVKPLGFIVSSIADGVTGTQVLVPAG 53


RESULT 11
C69297
acetyl-CoA decarbonylase/synthase, subunit beta (cdhC) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C69297
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: C69297
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-524 <KLE>
A;Cross-references: UNIPROT:O29868; UNIPARC:UPI0000057124; GB:AE001078; GB:AE000782; NI

  Query Match             32.6%;  Score 48.5;  DB 2;  Length 524;
  Best Local Similarity   50.0%;  Pred. No. 55;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          4 AQVKPLGPQLADPEVFLPQGELLD 27
              |:| | ||  |      :|:|||||
Db        246 AKVDPEGPNRA-----VPKGELLD 264


RESULT 12
D69117
nitrogenase iron-molybdenum cofactor biosynthesis protein NifB - Methanobacterium therm
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69117
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: D69117
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-288 <MTH>
A;Cross-references: UNIPROT:O27899; UNIPARC:UPI00000665ED; GB:AE000939; GB:AE000666; NI
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1871
```

```
  Query Match             32.2%;  Score 48;  DB 2;  Length 288;
  Best Local Similarity   37.9%;  Pred. No. 33;
  Matches   11;  Conservative    4;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              ||  :||   |  | :    :| ||: | |
Db        208 DIAREVKKRGASLMNIIPLIPMGEMKDYP 236


RESULT 13
A83702
inosine-uridine nucleoside hydrolase BH0417 [imported] - Bacillus halodurans (strain C-1
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A83702
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83702
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:Q9KFR1; UNIPARC:UPI00000C3888; GB:AP001508; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0417
C;Superfamily: yaaF protein

  Query Match             32.2%;  Score 48;  DB 2;  Length 309;
  Best Local Similarity   40.7%;  Pred. No. 36;
  Matches   11;  Conservative    4;  Mismatches   10;  Indels    2;  Gaps    1;

Qy          1 DILAQVKPL--GPQLADPEVFLPQGEL 25
              |:   : ||  |  :||   | |:|||
Db        260 DVSVSIAPLTRGQSVADFRPFAPEGEL 286


RESULT 14
T13993
hypothetical protein sf21 - common sunflower
C;Species: Helianthus annuus (common sunflower)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T13993
R;Kruter-Canham, R.; Bronner, R.; Evrand, J.L.; Hahne, G.; Friedt, W.; Steinmetz, A.
Plant Sci. 129, 191-202, 1997
A;Title: A transmitting tiussue-and pollen-expressed protein from sunflower with sequenc
A;Reference number: Z17848
A;Accession: T13993
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-352 <KRU>
A;Cross-references: UNIPROT:O23969; UNIPARC:UPI000013589F; EMBL:Y09057; NID:e1193468; PI
A;Experimental source: cell line HA401B; flower
C;Genetics:
A;Note: sf21

  Query Match             32.2%;  Score 48;  DB 2;  Length 352;
  Best Local Similarity   53.3%;  Pred. No. 42;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          8 PLGPQLADPEVFLPQ 22
              ||||   |||:  |:
Db        316 PLGPSSIDPELLYPE 330


RESULT 15
AC2251
hypothetical protein alr3562 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2251
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2251
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <KUR>
A;Cross-references: UNIPROT:Q8YR86; UNIPARC:UPI00000CE82C; GB:BA000019; PIDN:BAB75261.1
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3562

  Query Match             32.2%;  Score 48;  DB 2;  Length 473;
  Best Local Similarity   38.1%;  Pred. No. 58;
  Matches    8;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          9 LGPQLADPEVFLPQGELLDSP 29
              |  :|  | :::|| :||  |
Db        190 LTSKLTQPNIYVPQAQLLPQP 210


Search completed: March 21, 2006, 21:49:29
Job time : 20.5091 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:58 ; Search time 85.9455 Seconds
(without alignments)
238.062 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	100.0	529	1	CDSN_HUMAN	Q15517 homo sapien
2	149	100.0	529	1	CDSN_PANTR	Q7yr44 pan troglod
3	149	100.0	534	1	CDSN_MACMU	Q5tm45 macaca mula
4	99	66.4	561	1	CDSN_MOUSE	Q7tpc1 mus musculu
5	63	42.3	177	2	Q8H2M4_ORYSA	Q8h2m4 oryza sativ
6	59	39.6	558	2	Q93A36_ACILW	Q93a36 acinetobact
7	56	37.6	46	2	Q96JS6_HUMAN	Q96js6 homo sapien
8	56	37.6	300	1	ERA_MYCBO	P0a563 mycobacteri
9	56	37.6	300	1	ERA_MYCTU	P0a562 mycobacteri
10	55	36.9	264	2	Q4SBC0_TETNG	Q4sbc0 tetraodon n
11	54.5	36.6	381	2	Q4TKI8_9SPHN	Q4tki8 erythrobact
12	54	36.2	299	2	Q73Y13_MYCPA	Q73y13 mycobacteri
13	54	36.2	305	2	Q5YZW0_NOCFA	Q5yzw0 nocardia fa
14	54	36.2	1083	2	Q5QDM5_9BROM	Q5qdm5 fragaria ch
15	53	35.6	295	2	Q74M74_NANEQ	Q74m74 nanoarchaeu
16	53	35.6	312	2	Q9F3R2_9NOCA	Q9f3r2 rhodococcus
17	52.5	35.2	347	2	Q5N2Z6_SYNP6	Q5n2z6 synechococc
18	52	34.9	154	2	Q9YFA0_AERPE	Q9yfa0 aeropyrum p
19	52	34.9	234	2	Q4K8A8_PSEF5	Q4k8a8 pseudomonas
20	52	34.9	766	2	Q6ZMQ4_HUMAN	Q6zmq4 homo sapien
21	52	34.9	805	2	Q5B8K6_EMENI	Q5b8k6 aspergillus
22	52	34.9	939	2	Q967X6_DROME	Q967x6 drosophila
23	52	34.9	939	2	Q9VB35_DROME	Q9vb35 drosophila
24	51.5	34.6	246	2	Q8DEQ6_VIBVU	Q8deq6 vibrio vuln
25	51.5	34.6	256	2	Q8PWE0_METMA	Q8pwe0 methanosarc
26	51.5	34.6	264	2	Q7MNP9_VIBVY	Q7mnp9 vibrio vuln
27	51.5	34.6	381	2	Q6LG68_PHOPR	Q6lg68 photobacter
28	51.5	34.6	559	2	Q5AQN0_EMENI	Q5aqn0 aspergillus
29	51.5	34.6	837	1	RIX7_YEAST	Q07844 saccharomyc
30	51	34.2	100	2	Q9RJA6_STRCO	Q9rja6 streptomyce
31	51	34.2	265	2	Q98916_CHICK	Q98916 gallus gall
32	51	34.2	265	2	Q6DKC5_XENLA	Q6dkc5 xenopus lae
33	51	34.2	340	2	Q6A974_PROAC	Q6a974 propionibac
34	51	34.2	366	1	SERC_AZOSE	Q5p7u7 azoarcus sp
35	50.5	33.9	196	2	Q9XLB6_9TELE	Q9xlb6 marcusenius
36	50.5	33.9	207	2	Q6VF62_9NEOB	Q6vf62 heterixalus
37	50.5	33.9	211	1	FTHC_YEAST	P40099 saccharomyc
38	50.5	33.9	669	2	Q4WUC3_ASPFU	Q4wuc3 aspergillus
39	50.5	33.9	1203	2	Q8BXP6_MOUSE	Q8bxp6 mus musculu
40	50	33.6	98	2	Q76AW8_9CALI	Q76aw8 human norov
41	50	33.6	258	2	Q65DE7_BACLD	Q65de7 bacillus li
42	50	33.6	280	2	Q7NL30_GLOVI	Q7nl30 gloeobacter
43	50	33.6	289	2	Q8GGT6_STRAZ	Q8ggt6 streptomyce
44	50	33.6	323	2	Q609B4_METCA	Q609b4 methylococc
45	50	33.6	339	2	Q8M5W2_9CHAR	Q8m5w2 limnodromus

ALIGNMENTS

RESULT 1
CDSN_HUMAN
ID CDSN_HUMAN STANDARD; PRT; 529 AA.
AC Q15517; O43509; Q5SQ85; Q5STD2; Q7LA70; Q7LA71; Q86Z04; Q8IZU4;
AC Q8IZU5; Q8IZU6; Q8N5P3; Q95IF9; Q9NP52; Q9NPE0; Q9NPG5; Q9NRH4;
AC Q9NRH5; Q9NRH6; Q9NRH7; Q9NRH8; Q9UBH8; Q9UIN6; Q9UIN7; Q9UIN8;
AC Q9UIN9; Q9UIP0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Corneodesmosin precursor (S protein).
GN Name=CDSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT LEU-18.
RC TISSUE=Foreskin;
RX MEDLINE=94022396; PubMed=8415725;
RA Zhou Y., Chaplin D.D.;
RT "Identification in the HLA class I region of a gene expressed late in
RT keratinocyte differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9470-9474(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT ALA-408.
RC TISSUE=Epidermis;
RX MEDLINE=98380490; PubMed=9712893; DOI=10.1074/jbc.273.35.22640;
RA Guerrin M., Simon M., Montezin M., Haftek M., Vincent C., Serre G.;
RT "Expression cloning of human corneodesmosin proves its identity with
RT the product of the S gene and allows improved characterization of its
RT processing during keratinocyte differentiation.";
RL J. Biol. Chem. 273:22640-22647(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS LEU-18; SER-153 DEL;
RP PHE-202 AND LEU-410.
RX MEDLINE=22256218; PubMed=12366786;
RX DOI=10.1034/j.1399-0039.2002.600110.x;
RA Hui J., Oka A., Tamiya G., Tomizawa M., Kulski J.K., Penhale W.J.,
RA Tay G.K., Iizuka M., Ozawa A., Inoko H.;
RT "Corneodesmosin DNA polymorphisms in MHC haplotypes and Japanese
RT patients with psoriasis.";
RL Tissue Antigens 60:77-83(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Shiina T., Ota M., Katsuyama Y., Hashimoto N., Inoko H.;
RT "Genome diversity in HLA: a new strategy for detection of genetic
RT polymorphisms in expressed genes within the HLA class III and class I

RT regions.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND VARIANTS LEU-18;
RP ASN-143; SER-153 DEL; PHE-202; ALA-408; LEU-410 AND ASN-527.
RX MEDLINE=22935763; PubMed=14574404; DOI=10.1038/nature02055;
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA Gilbert J.G.R., Clamp M.E., Bethel G., Milne S., Ainscough R.,
RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,
RA Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA Blakey S.E., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcott R.,
RA Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLay K.,
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA Porter K.M., Ramsey Y., Ranby S.A., Rice C.M., Ross M.T., Searle S.M.,
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.L.,
RA Whittaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT "The DNA sequence and analysis of human chromosome 6.";
RL Nature 425:805-811(2003).
RN [7]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANTS LEU-18; SER-153
RP DEL; PHE-202 AND LEU-410.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 30-529 (ALLELES 1.11; 1.21; 1.31;
RP 1.32; 1.41; 1.42; 1.43; 1.51; 1.52; 2.11; 2.21; 2.22 AND 2.23), AND
RP VARIANTS PHE-56; ASN-143 DEL; ASN-150; PHE-202; GLY-401; ALA-408;
RP LEU-410 AND ASN-527.
RX MEDLINE=21108612; PubMed=11169256;
RX DOI=10.1034/j.1399-0039.2001.057001032.x;
RA Guerrin M., Vincent C., Simon M., Ahnini R.T., Fort M., Serre G.;
RT "Identification of six novel polymorphisms in the human corneodesmosin
RT gene.";
RL Tissue Antigens 57:32-38(2001).
RN [9]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 106-529, VARIANTS ASN-143;
RP PHE-202; GLY-401; ALA-408 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX MEDLINE=20066866; PubMed=10599883;
RX DOI=10.1034/j.1399-0039.1999.540501.x;
RA Jenisch S., Koch S., Henseler T., Nair R.P., Elder J.T., Watts C.E.,
RA Westphal E., Voorhees J.J., Christophers E., Kroenke M.;
RT "Corneodesmosin gene polymorphism demonstrates strong linkage
RT disequilibrium with HLA and association with psoriasis vulgaris.";
RL Tissue Antigens 54:439-449(1999).
RN [10]
RP PARTIAL PROTEIN SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=98058975; PubMed=9395522; DOI=10.1074/jbc.272.50.31770;
RA Simon M., Montezin M., Guerrin M., Durieux J.-J., Serre G.;
RT "Characterization and purification of human corneodesmosin, an
RT epidermal basic glycoprotein associated with corneocyte-specific
RT modified desmosomes.";
RL J. Biol. Chem. 272:31770-31776(1997).
RN [11]
RP INVOLVEMENT IN HTSS.
RX PubMed=12754508; DOI=10.1038/ng1163;
RA Levy-Nissenbaum E., Betz R.C., Frydman M., Simon M., Lahat H.,
RA Bakhan T., Goldman B., Bygum A., Pierick M., Hillmer A.M., Jonca N.,
RA Toribio J., Kruse R., Dewald G., Cichon S., Kubisch C., Guerrin M.,
RA Serre G., Nothen M.M., Pras E.;
RT "Hypotrichosis simplex of the scalp is associated with nonsense
RT mutations in CDSN encoding corneodesmosin.";
RL Nat. Genet. 34:151-153(2003).
RN [12]
RP VARIANTS PHE-202 AND LEU-410, AND ASSOCIATION WITH PSORIASIS.
RX PubMed=10844560; DOI=10.1046/j.1523-1747.2000.00002.x;
RA Enerbaeck C., Enlund F., Inerot A., Samuelsson L., Wahlstroem J.,
RA Swanbeck G., Martinsson T.;
RT "S gene (Corneodesmosin) diversity and its relationship to psoriasis;
RT high content of cSNP in the HLA-linked S gene.";
RL J. Invest. Dermatol. 114:1158-1163(2000).
RN [13]
RP VARIANTS PHE-202; GLY-401; ALA-408; LEU-410 AND ASN-527, AND
RP ASSOCIATION WITH PSORIASIS.
RX PubMed=12472658; DOI=10.1034/j.1399-0039.2002.600403.x;
RA Orru S., Giuressi E., Casula M., Loizedda A., Murru R., Mulargia M.,
RA Masala M.V., Cerimele D., Zucca M., Aste N., Biggio P., Carcassi C.,
RA Contu L.;
RT "Psoriasis is associated with a SNP haplotype of the corneodesmosin
RT gene (CDSN).";
RL Tissue Antigens 60:292-298(2002).
CC -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC intercellular structures that are involved in desquamation.
CC -!- TISSUE SPECIFICITY: Exclusively expressed in skin.
CC -!- POLYMORPHISM: Genetic variation in CDSN may be associated with
CC susceptibility to psoriasis [MIM:177900]. Various CDSN alleles are
CC known including alleles 1.11, 1.21, 1.31, 1.32, 1.41, 1.42, 1.43,
CC 1.51, 1.52, 2.11, 2.21, 2.22 and 2.23.
CC -!- DISEASE: Defects in CDSN are a cause of hypotrichosis simplex of
CC the scalp (HTSS) [MIM:146520]; also known as hypotrichosis Spanish
CC type. HTSS is an autosomal dominant form of isolated alopecia.
CC Affected individuals have normal hair in early childhood but
CC experience progressive loss of scalp hair beginning in the middle
CC of the first decade and almost complete baldness by the third
CC decade.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 501.
CC ---
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ---

```
DR   EMBL; L20815; AAA21321.1; ALT_SEQ; mRNA.
DR   EMBL; AF030130; AAC24196.1; -; mRNA.

  Query Match             100.0%;  Score 149;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 2
CDSN_PANTR
ID   CDSN_PANTR     STANDARD;      PRT;   529 AA.
AC   Q7YR44;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor (S protein).
GN   Name=CDSN;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=22709134; PubMed=12799463; DOI=10.1073/pnas.1230533100;
RA   Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
RA   Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA   Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA   Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT   "Comparative sequencing of human and chimpanzee MHC class I regions
RT   unveils insertions/deletions as the major path to genomic
RT   divergence.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BA000041; BAC78170.1; -; Genomic_DNA.
KW   Glycoprotein; Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    529       Corneodesmosin.
FT   COMPBIAS     62    464       Ser-rich.
FT   CARBOHYD    172    172       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   529 AA;  51611 MW;  7E7111EC28D0B018 CRC64;

  Query Match             100.0%;  Score 149;  DB 1;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 3
CDSN_MACMU
ID   CDSN_MACMU     STANDARD;      PRT;   534 AA.
AC   Q5TM45;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=CDSN;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   PubMed=15269276; DOI=10.1093/molbev/msh216;
RA   Kulski J.K., Anzai T., Shiina T., Inoko H.;
RT   "Rhesus macaque class I duplicon structures, organization, and
RT   evolution within the alpha block of the major histocompatibility
RT   complex.";
RL   Mol. Biol. Evol. 21:2079-2091(2004).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB128049; BAD69749.1; -; Genomic_DNA.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    534       Corneodesmosin.
FT   COMPBIAS     62    469       Ser-rich.
SQ   SEQUENCE   534 AA;  51983 MW;  CB6A9DF256C32874 CRC64;

  Query Match             100.0%;  Score 149;  DB 1;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        506 DILAQVKPLGPQLADPEVFLPQGELLDSP 534


RESULT 4
CDSN_MOUSE
ID   CDSN_MOUSE     STANDARD;      PRT;   561 AA.
AC   Q7TPC1;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Corneodesmosin precursor.
GN   Name=Cdsn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C3H/He; TISSUE=Osteoblast;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

```
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Secreted; found in corneodesmosomes; the
CC       intercellular structures that are involved in desquamation (By
CC       similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; BC055373; AAH55373.1; -; mRNA.
DR   Ensembl; ENSMUSG00000039518; Mus musculus.
DR   MGI; MGI:3505689; Cdsn.
KW   Signal.
FT   SIGNAL        1     32       Potential.
FT   CHAIN        33    561       Corneodesmosin.
FT   COMPBIAS     62    493       Ser-rich.
SQ   SEQUENCE   561 AA;  54414 MW;  7A6EBEF364038746 CRC64;

  Query Match             66.4%;  Score 99;  DB 1;  Length 561;
  Best Local Similarity   69.0%;  Pred. No. 2.6e-05;
  Matches   20;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              :|| ||||||||| ||:| |||||    |
Db        529 NILTQVKPLGPQLMDPKVSLPQGEPQGEP 557


RESULT 5
Q8H2M4_ORYSA
ID   Q8H2M4_ORYSA PRELIMINARY;      PRT;   177 AA.
AC   Q8H2M4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein OSJNBa0066H10.129.
GN   Name=OSJNBa0066H10.129;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OSJNBa0066H10.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005516; BAC21595.1; -; Genomic_DNA.
DR   Gramene; Q8H2M4; -.
KW   Hypothetical protein.
SQ   SEQUENCE   177 AA;  20173 MW;  C24FDC156AF093FB CRC64;

  Query Match             42.3%;  Score 63;  DB 2;  Length 177;
  Best Local Similarity   50.0%;  Pred. No. 1.1;
  Matches   13;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELL 26
              |: ||:: |  ||||  || : | |
Db         24 DLAAQLRALAAQLADPTAFLEEAEPL 49


RESULT 6
Q93A36_ACILW
ID   Q93A36_ACILW PRELIMINARY;      PRT;   558 AA.
AC   Q93A36;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Type IIs restriction endonuclease R.AlwI.
OS   Acinetobacter lwoffii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Moraxellaceae; Acinetobacter.
OX   NCBI_TaxID=28090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Lunnen K.D., Wilson G.G.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF431889; AAL25126.1; -; Genomic_DNA.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
KW   Endonuclease.
SQ   SEQUENCE   558 AA;  63894 MW;  0F2D3D8B671A865B CRC64;

  Query Match             39.6%;  Score 59;  DB 2;  Length 558;
  Best Local Similarity   41.7%;  Pred. No. 14;
  Matches   10;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGE 24
              |:|   ||:  ::|| |: :| ||
Db        350 DLLISPKPISIEIADKEISIPSGE 373


RESULT 7
Q96JS6_HUMAN
ID   Q96JS6_HUMAN PRELIMINARY;      PRT;    46 AA.
AC   Q96JS6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC008248; AAH08248.1; -; mRNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   46 AA;  4663 MW;  09660CABFB0F763D CRC64;

  Query Match             37.6%;  Score 56;  DB 2;  Length 46;
  Best Local Similarity   45.5%;  Pred. No. 2.6;
  Matches   10;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          8 PLGPQLADPEVFLPQGELLDSP 29
              | || |  |:: || |:|  :|
Db         10 PAGPGLCSPQLLLPGGQLFWAP 31


RESULT 8
ERA_MYCBO
ID   ERA_MYCBO      STANDARD;      PRT;   300 AA.
AC   P0A563; O05834;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   GTP-binding protein era homolog.
GN   Name=era; OrderedLocusNames=Mb2385c;
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1765;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
```

```
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC   -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
CC       and is essential for cell growth (By similarity).
CC   -!- SIMILARITY: Belongs to the era/trmE GTP-binding protein family.
CC       Era subfamily.
CC   -!- SIMILARITY: Contains 1 KH type-2 domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BX248342; CAD97246.1; -; Genomic_DNA.
DR   HSSP; Q9X1F8; 1MKY.
DR   HAMAP; MF_00367; -; 1.
DR   InterPro; IPR005662; Era_GTP_bd.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR004044; KH_TYPE_2.
DR   InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR   InterPro; IPR005225; Small_GTP_bd.
DR   PANTHER; PTHR11649:SF4; Era; 1.
DR   Pfam; PF07650; KH_2; 1.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   TIGRFAMs; TIGR00436; era; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50823; KH_TYPE_2; 1.
KW   Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
FT   DOMAIN      207    286       KH type-2.
FT   NP_BIND      13     20       GTP (Potential).
FT   NP_BIND      60     64       GTP (Potential).
FT   NP_BIND     125    128       GTP (Potential).
SQ   SEQUENCE   300 AA;  32460 MW;  01A5C664C5E9D9D7 CRC64;

  Query Match             37.6%;  Score 56;  DB 1;  Length 300;
  Best Local Similarity   41.4%;  Pred. No. 20;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        168 DVLAAALPAGP------AYYPDGELTDEP 190


RESULT 9
ERA_MYCTU
ID   ERA_MYCTU      STANDARD;      PRT;   300 AA.
AC   P0A562; O05834;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   GTP-binding protein era homolog.
GN   Name=era; OrderedLocusNames=Rv2364c, MT2433; ORFNames=MTCY27.16;
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1773;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA   Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA   Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA   Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA   Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA   Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=CDC 1551 / Oshkosh;
RX   MEDLINE=22206494; PubMed=12218036;
RX   DOI=10.1128/JB.184.19.5479-5490.2002;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA   Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA   Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA   Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA   Fraser C.M.;
RT   "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   J. Bacteriol. 184:5479-5490(2002).
CC   -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
CC       and is essential for cell growth (By similarity).
CC   -!- SIMILARITY: Belongs to the era/trmE GTP-binding protein family.
CC       Era subfamily.
CC   -!- SIMILARITY: Contains 1 KH type-2 domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BX842579; CAE55479.1; -; Genomic_DNA.
DR   EMBL; AE000516; AAK46727.1; -; Genomic_DNA.
DR   PIR; C70586; C70586.
DR   HSSP; Q9X1F8; 1MKY.
DR   TIGR; MT2433; -.
DR   TubercuList; Rv2364c; -.
DR   HAMAP; MF_00367; -; 1.
DR   InterPro; IPR005662; Era_GTP_bd.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR004044; KH_TYPE_2.
DR   InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR   InterPro; IPR005225; Small_GTP_bd.
DR   PANTHER; PTHR11649:SF4; Era; 1.
DR   Pfam; PF07650; KH_2; 1.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   TIGRFAMs; TIGR00436; era; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50823; KH_TYPE_2; 1.
KW   Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
FT   DOMAIN      207    286       KH type-2.
FT   NP_BIND      13     20       GTP (Potential).
FT   NP_BIND      60     64       GTP (Potential).
FT   NP_BIND     125    128       GTP (Potential).
SQ   SEQUENCE   300 AA;  32460 MW;  01A5C664C5E9D9D7 CRC64;

  Query Match             37.6%;  Score 56;  DB 1;  Length 300;
  Best Local Similarity   41.4%;  Pred. No. 20;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        168 DVLAAALPAGP------AYYPDGELTDEP 190


RESULT 10
Q4SBC0_TETNG
ID   Q4SBC0_TETNG PRELIMINARY;      PRT;   264 AA.
AC   Q4SBC0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
```

```
DE   Chromosome undetermined SCAF14676, whole genome shotgun sequence.
GN   ORFNames=GSTENG00021041001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014676; CAG02062.1; -; Genomic_DNA.
SQ   SEQUENCE   264 AA;  31351 MW;  A91843C8331F05E1 CRC64;

  Query Match             36.9%;  Score 55;  DB 2;  Length 264;
  Best Local Similarity   52.4%;  Pred. No. 24;
  Matches   11;  Conservative    5;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          7 KPLGPQLADPEVFLPQGELLD 27
              ||: |||| |::  |:|| :|
Db         46 KPMVPQLAPPKI--PEGERVD 64


RESULT 11
Q4TKI8_9SPHN
ID   Q4TKI8_9SPHN PRELIMINARY;      PRT;   381 AA.
AC   Q4TKI8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Glycosyltransferase.
GN   ORFNames=ELI2639;
OS   Erythrobacter litoralis HTCC2594.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Erythrobacter.
OX   NCBI_TaxID=314225;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HTCC2594;
RA   Giovannoni S.J., Cho J.-C., Ferriera S., Johnson J., Kravitz S.,
RA   Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
RA   Friedman R., Venter J.C.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAGG01000008; EAL74312.1; -; Genomic_DNA.
KW   Transferase.
SQ   SEQUENCE   381 AA;  42084 MW;  A85CDD1FA5C53C3C CRC64;

  Query Match             36.6%;  Score 54.5;  DB 2;  Length 381;
  Best Local Similarity   46.2%;  Pred. No. 42;
  Matches   12;  Conservative    4;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          4 AQVKPLGPQLADPEVFLPQGELLDSP 29
                |:|   |:: ||: | | ||| | |
Db         38 AAVRVFSPKVRDPD-FEPTGELFDVP 62


RESULT 12
Q73Y13_MYCPA
ID   Q73Y13_MYCPA PRELIMINARY;      PRT;   299 AA.
AC   Q73Y13;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Bex.
GN   Name=bex; OrderedLocusNames=MAP2145c;
OS   Mycobacterium paratuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium avium complex (MAC).
OX   NCBI_TaxID=1770;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=k10;
RA   Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017234; AAS04462.1; -; Genomic_DNA.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   InterPro; IPR005662; Era_GTP_bd.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR004044; KH_TYPE_2.
DR   InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR   InterPro; IPR005225; Small_GTP_bd.
DR   Pfam; PF07650; KH_2; 1.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   TIGRFAMs; TIGR00436; era; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50823; KH_TYPE_2; 1.
KW   Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
SQ   SEQUENCE   299 AA;  32548 MW;  16CA66E229DF91C2 CRC64;

  Query Match             36.2%;  Score 54;  DB 2;  Length 299;
  Best Local Similarity   41.4%;  Pred. No. 38;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        167 DVLAAALPPGP------AYYPDGELTDEP 189


RESULT 13
Q5YZW0_NOCFA
ID   Q5YZW0_NOCFA PRELIMINARY;      PRT;   305 AA.
AC   Q5YZW0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Putative GTP-binding protein.
GN   OrderedLocusNames=nfa14360;
OS   Nocardia farcinica.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Nocardiaceae; Nocardia.
OX   NCBI_TaxID=37329;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IFM 10152;
RX   PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA   Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA   Shiba T., Hattori M.;
RT   "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
```

```
CC   -!- SIMILARITY: Belongs to the era/trmE GTP-binding protein family.
DR   EMBL; AP006618; BAD56281.1; -; Genomic_DNA.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   InterPro; IPR005662; Era_GTP_bd.
DR   InterPro; IPR005289; GTP_bd.
DR   InterPro; IPR004044; KH_TYPE_2.
DR   InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR   InterPro; IPR005225; Small_GTP_bd.
DR   Pfam; PF07650; KH_2; 1.
DR   Pfam; PF01926; MMR_HSR1; 1.
DR   TIGRFAMs; TIGR00436; era; 1.
DR   TIGRFAMs; TIGR00650; MG442; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS50823; KH_TYPE_2; 1.
KW   Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
SQ   SEQUENCE   305 AA;  33520 MW;  35CF8232762F167B CRC64;

  Query Match             36.2%;  Score 54;  DB 2;  Length 305;
  Best Local Similarity   41.4%;  Pred. No. 39;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |::|   | ||       | | ||| | |
Db        173 DVIAAKMPEGP------AFYPDGELTDEP 195


RESULT 14
Q5QDM5_9BROM
ID   Q5QDM5_9BROM PRELIMINARY;      PRT;  1083 AA.
AC   Q5QDM5;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Replicase.
OS   Fragaria chiloensis latent virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC   Ilarvirus.
OX   NCBI_TaxID=255238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15664053; DOI=10.1016/j.jviromet.2004.11.006;
RA   Tzanetakis I.E., Keller K.E., Martin R.R.;
RT   "The use of reverse transcriptase for efficient first- and second-
RT   strand cDNA synthesis from single- and double-stranded RNA
RT   templates.";
RL   J. Virol. Methods 124:73-77(2005).
DR   EMBL; AY682102; AAV68306.1; -; Genomic_RNA.
DR   GO; GO:0008174; F:mRNA methyltransferase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003724; F:RNA helicase activity; IEA.
DR   GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR   GO; GO:0006396; P:RNA processing; IEA.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR002588; V_methyltrans.
DR   InterPro; IPR000606; Viral_helicase1.
DR   Pfam; PF01443; Viral_helicase1; 1.
DR   Pfam; PF01660; Vmethyltransf; 1.
SQ   SEQUENCE   1083 AA;  121136 MW;  74C05DA7D3E3DB8E CRC64;

  Query Match             36.2%;  Score 54;  DB 2;  Length 1083;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ILAQVKPLGPQLADPEVFLP 21
              || : : : | |||||:|:|
Db        521 ILMEYQVMTPYLADPELFVP 540


RESULT 15
Q74M74_NANEQ
ID   Q74M74_NANEQ PRELIMINARY;      PRT;   295 AA.
AC   Q74M74;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   NEQ440.
GN   OrderedLocusNames=NEQ440;
OS   Nanoarchaeum equitans.
OC   Archaea; Nanoarchaeota; Nanoarchaeum.
OX   NCBI_TaxID=160232;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Kin4-M;
RX   MEDLINE=22946215; PubMed=14566062; DOI=10.1073/pnas.1735403100;
RA   Waters E., Hohn M.J., Ahel I., Graham D.E., Adams M.D., Barnstead M.,
RA   Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X.,
RA   Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
RA   Soell D., Stetter K.O., Short J.M., Noorderwier M.;
RT   "The genome of Nanoarchaeum equitans: insights into early archaeal
RT   evolution and derived parasitism.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988(2003).
DR   EMBL; AE017199; AAR39284.1; -; Genomic_DNA.
DR   GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000241; RNA_methylase.
DR   InterPro; IPR000051; SAM_bd.
DR   Pfam; PF01170; UPF0020; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW   Complete proteome; Methyltransferase; Transferase.
SQ   SEQUENCE   295 AA;  34066 MW;  1BB1499891200891 CRC64;

  Query Match             35.6%;  Score 53;  DB 2;  Length 295;
  Best Local Similarity   46.2%;  Pred. No. 53;
  Matches   12;  Conservative    5;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          6 VKPLGPQLADPEVFLPQGE--LLDSP 29
              |  | |:  :||:|:|:|   ||| |
Db        250 VSMLYPEFLNPELFVPKGYKILLDRP 275


Search completed: March 21, 2006, 21:48:47
Job time : 88.9455 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:42:43 ; Search time 93.3273 Seconds
(without alignments)
136.530 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	100.0	58	7	ABJ39194	Abj39194 Fragment
2	149	100.0	529	2	AAW80498	Aaw80498 A protein
3	149	100.0	529	2	AAW85055	Aaw85055 Human pro
4	149	100.0	529	3	AAY44989	Aay44989 Human epi
5	149	100.0	529	4	AAU27550	Aau27550 Corneodes
6	149	100.0	529	7	ABJ39148	Abj39148 Human cor
7	149	100.0	529	8	ADN04226	Adn04226 Antipsori
8	149	100.0	529	8	ADP12525	Adp12525 Protein e
9	142	95.3	529	4	AAG64152	Aag64152 Polypepti
10	61.5	41.3	704	9	ADW17250	Adw17250 Eucalyptu
11	59	39.6	252	4	ABB06906	Abb06906 Micromono
12	59	39.6	252	6	ABP99287	Abp99287 Orthosomy
13	56	37.6	300	6	ABU36775	Abu36775 Protein e
14	56	37.6	309	6	ABU34700	Abu34700 Protein e
15	55	36.9	272	3	AAY83764	Aay83764 Encephali
16	54.5	36.6	410	6	ADB12079	Adb12079 Alloiococ
17	54.5	36.6	428	6	ADB12077	Adb12077 Alloiococ
18	54.5	36.6	433	6	ADB12075	Adb12075 Alloiococ
19	54	36.2	299	6	ABU34258	Abu34258 Protein e
20	53	35.6	295	8	ADK16907	Adk16907 Nanoarcha
21	52	34.9	312	4	ABG04075	Abg04075 Novel hum
22	52	34.9	766	8	ADR50797	Adr50797 Human c-b
23	51.5	34.6	837	6	ABR53780	Abr53780 Protein s
24	51.5	34.6	837	7	ADK63834	Adk63834 Disease t
25	51.5	34.6	837	8	ADS43866	Ads43866 Bacterial
26	51	34.2	140	5	ABP04147	Abp04147 Human ORF
27	51	34.2	308	4	AAU47020	Aau47020 Propionib
28	51	34.2	308	6	ABM43539	Abm43539 Propionib
29	51	34.2	353	6	ABM64792	Abm64792 Propionib
30	51	34.2	1069	8	ADR38465	Adr38465 Cellulosi
31	51	34.2	1107	8	ADR38464	Adr38464 Cellulosi
32	51	34.2	1168	4	AAM79432	Aam79432 Human pro
33	51	34.2	1168	4	AAM40573	Aam40573 Human pol
34	51	34.2	1169	4	AAM38787	Aam38787 Human pol
35	51	34.2	4561	4	ABG30203	Abg30203 Novel hum
36	51	34.2	9222	4	ABG21064	Abg21064 Novel hum
37	50	33.6	283	8	ADY09424	Ady09424 Plant ful
38	50	33.6	289	6	ABU11341	Abu11341 Protein e
39	50	33.6	353	4	ABG07351	Abg07351 Novel hum
40	50	33.6	353	7	ADE09312	Ade09312 Novel pro
41	50	33.6	355	4	ABG18948	Abg18948 Novel hum
42	50	33.6	400	8	ADS42217	Ads42217 Bacterial
43	50	33.6	402	7	ADE08349	Ade08349 Novel pro
44	50	33.6	1440	5	ABB54801	Abb54801 Lactococc
45	49.5	33.2	277	5	ABB82142	Abb82142 B. cereus

ALIGNMENTS

RESULT 1
ABJ39194
ID ABJ39194 standard; protein; 58 AA.
XX
AC ABJ39194;
XX
DT 06-NOV-2003 (first entry)
XX
DE Fragment b, cleavage product of human S (corneodesmosin skin) protein.
XX
KW Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW winter xerosis; ichthyoses; dermatological; gene therapy; S protein;
KW skin.
XX
OS Homo sapiens.
XX
PN WO2003050311-A2.
XX
PD 19-JUN-2003.
XX
PF 09-DEC-2002; 2002WO-US039485.
XX
PR 07-DEC-2001; 2001US-0338188P.
XX
PA (WELA) WELLA AG.
XX
PI Pras E, Levy-Nissenbaum E, Frydman M, Thederahn T;
XX
DR WPI; 2003-569100/53.
XX
PT Therapeutic composition comprising an effector of a corneodesmosin gene
PT function associated with hair growth and a carrier, useful for treating
PT alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS Disclosure; Fig 2e; 37pp; English.
XX
CC This invention relates to the identification of genes linked to hair loss
CC and the retardation of hair growth. Alopecia is the general term used to
CC refer to loss of hair, the causes of which can include autoimmune
CC disorders, hormonal imbalances, neurological conditions and genetic
CC abnormalities. Specifically, the invention describes the highly conserved
CC corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC in the onset of nonsyndromic alopecia, an example being hytrichosis
CC simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases

```
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   peptide sequence is a functional region of the human S protein, the
CC   cornesodesmosin protein solely expressed in the skin that is a target of
CC   the invention
XX
SQ   Sequence 58 AA;

  Query Match             100.0%;  Score 149;  DB 7;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-14;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db         30 DILAQVKPLGPQLADPEVFLPQGELLDSP 58


RESULT 2
AAW80498
ID   AAW80498 standard; protein; 529 AA.
XX
AC   AAW80498;
XX
DT   10-FEB-1999  (first entry)
XX
DE   A protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761362-A1.
XX
PD   02-OCT-1998.
XX
PF   28-MAR-1997;   97FR-00003899.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586093/50.
DR   N-PSDB; AAV63614.
XX
PT   New polypeptide, corneodesmosin, and related nucleic acid - for cosmetic
PT   or therapeutic treatment of trophic disorders or thinning of the
PT   epidermis.
XX
PS   Claim 1; Page 29; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 149;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 3
AAW85055
ID   AAW85055 standard; protein; 529 AA.
XX
AC   AAW85055;
XX
DT   05-FEB-1999  (first entry)
XX
DE   Human protein designated corneodesmosin.
XX
KW   Corneodesmosin; intercorneocyte cohesion; skin trophic disorder;
KW   scar formation; epidermis thinning; skin fragility; hyperkeratosis;
KW   xerosis; ichthyosis; psoriasis; hyperkeratotic tumour;
KW   reactive keratosis; and leucokeratosis.
XX
OS   Homo sapiens.
XX
PN   FR2761363-A1.
XX
PD   02-OCT-1998.
XX
PF   11-SEP-1997;   97FR-00011317.
XX
PR   28-MAR-1997;   97FR-00003899.
XX
PA   (OREA ) L'OREAL SA.
XX
PI   Serre G,  Simon M,  Weber VM;
XX
DR   WPI; 1998-586094/50.
DR   N-PSDB; AAV63541.
XX
PT   New polypeptide, corneodesmosin, from skin and related nucleic acid - for
PT   cosmetic or therapeutic treatment of trophic disorders, thinning of the
PT   epidermis etc., also compositions containing protease that degrades
PT   corneodesmosin.
XX
PS   Claim 1; Page 30; 35pp; French.
XX
CC   The present sequence represents a protein designated corneodesmosin. The
CC   protein is involved in intercorneocyte cohesion and is used, in cosmetic
CC   or pharmaceutical compositions, to treat trophic disorders of the skin
CC   and disorders associated with scar formation, and thinning of the
CC   epidermis, particularly the corneal layer, or skin fragility, to
CC   reinforce intercorneocyte cohesion and/or to induce thickening of the
CC   corneal layer. Hyperkeratosis, xerosis, ichthyosis, psoriasis, malignant
CC   or benign hyperkeratotic tumours (including those of the Malphigian
CC   mucosa), reactive keratosis, and leucokeratosis (of the cervix uteri,
CC   mouth or during prolapse) can also be treated with the products
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 149;  DB 2;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 4
AAY44989
ID   AAY44989 standard; protein; 529 AA.
XX
AC   AAY44989;
XX
```

```
DT   23-MAY-2000  (first entry)
XX
DE   Human epidermal protein-6.
XX
KW   Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW   dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW   arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .32
FT                   /label= Signal_peptide
FT   Modified-site   3
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   37
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   50
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   57
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   80
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   105
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   172
FT                   /note= "N-glycosylated"
FT   Modified-site   182
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   287
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   395
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   440
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
FT   Modified-site   498
FT                   /note= "Potential phosphorylation site"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017107.
XX
PR   28-JUL-1998;   98US-0155203P.
PR   07-DEC-1998;   98US-0155254P.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   N-PSDB; AAZ50582.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders.
XX
PS   Claim 1; Fig 4; 82pp; English.
XX
CC   The present sequence is human epidermal protein-6 (HEPI) expressed in
CC   breast skin and keratinocytes. This protein is derived from PTHYNOT03
CC   cDNA library. HEPI-6 shares 99% identity with human S protein.
CC   Recombinant vectors comprising HEPI cDNA are introduced into host cells
CC   for protein expression. The HEPI proteins are useful for the treatment of
CC   epithelial disorders, including dyshidrotic eczema and scabies, cell
CC   proliferative disorders including actinic keratosis and arteriosclerosis,
CC   and autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 149;  DB 3;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 5
AAU27550
ID   AAU27550 standard; protein; 529 AA.
XX
AC   AAU27550;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Corneodesmosin.
XX
KW   Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW   antipsoriatic; corneodesmosin; inflammatory disease; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200162788-A2.
XX
PD   30-AUG-2001.
XX
PF   23-FEB-2001; 2001WO-GB000795.
XX
PR   23-FEB-2000; 2000GB-00004312.
XX
PA   (OXAG-) OXAGEN LTD.
XX
PI   Olaveson M,  Lench N,  Allen M,  Tazi-Ahnini R;
XX
DR   WPI; 2001-570627/64.
DR   N-PSDB; AAS43749.
XX
PT   Corneodesmosin protein and polynucleotide encoding it, having one or more
PT   polymorphisms useful in treating, diagnosing or determining
PT   susceptibility to corneodesmosin-mediated diseases, for e.g. inflammatory
PT   diseases.
XX
PS   Claim 8; Fig 2; 60pp; English.
XX
CC   The invention relates to corneodesmosin protein (I) and nucleic acid (II)
CC   encoding the corneodesmosin gene, where the gene comprises a base
CC   substitution, deletion or insertion at one or more positions. (I) and
CC   (II) are useful for screening for agents for use in prognosis, diagnosis
CC   and treatment of individuals having or being susceptible to
CC   corneodesmosin-mediated disease, by monitoring the reaction between the
CC   molecules and the agents. The nucleotide and amino acid polymorphisms are
CC   useful for diagnosing or determining susceptible to corneodesmosin-
CC   mediated disease, which facilitates subsequent treatment of the disease
CC   for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC   are useful in diagnostic, prognostic or therapeutic methods and as
CC   research tools for e.g. in drug screening. (II) is useful as probes or
CC   primers for detecting an allele of the polymorphism or in the regulation
CC   of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC   screening DNA clone libraries for cells secreting the antigen. (II) is
CC   useful as a model to investigate the role of corneodesmosin in normal
CC   skin function. The present sequence represents corneodesmosin amino acid
CC   sequence, as described in the invention
XX
SQ   Sequence 529 AA;
```

```
  Query Match            100.0%;  Score 149;  DB 4;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 6
ABJ39148
ID   ABJ39148 standard; protein; 529 AA.
XX
AC   ABJ39148;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human corneodesmosin (CDSN) protein.
XX
KW   Human; corneodesmosin; CDSN; hair loss; autoimmune disorder; psoriasis;
KW   neurological; hormonal; nonsyndromic alopecia; male pattern baldness;
KW   hytrichosis simplex of the scalp; HSS; hyperkeratotic cutaneous disease;
KW   winter xerosis; ichthyoses; dermatological; gene therapy; chromosome 6.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 28
FT                   /note= "Encoded by CCAG"
FT   misc_difference 408
FT                   /note= "Encoded by TCC"
XX
PN   WO2003050311-A2.
XX
PD   19-JUN-2003.
XX
PF   09-DEC-2002; 2002WO-US039485.
XX
PR   07-DEC-2001; 2001US-0338188P.
XX
PA   (WELA ) WELLA AG.
XX
PI   Pras E,  Levy-Nissenbaum E,  Frydman M,  Thederahn T;
XX
DR   WPI; 2003-569100/53.
DR   N-PSDB; ABT43849, ABT43850.
XX
PT   Therapeutic composition comprising an effector of a corneodesmosin gene
PT   function associated with hair growth and a carrier, useful for treating
PT   alopecia, particularly hytrichosis simplex of the scalp (HSS).
XX
PS   Disclosure; Fig 2e; 37pp; English.
XX
CC   This invention relates to the identification of genes linked to hair loss
CC   and the retardation of hair growth. Alopecia is the general term used to
CC   refer to loss of hair, the causes of which can include autoimmune
CC   disorders, hormonal imbalances, neurological conditions and genetic
CC   abnormalities. Specifically, the invention describes the highly conserved
CC   corneodesmosin (CDSN) gene and encoded protein thereof that play a role
CC   in the onset of nonsyndromic alopecia, an example being hytrichosis
CC   simplex of the scalp (HSS). Furthermore, human CDSN (localised to
CC   chromosome 6) is overexpressed in many hyperkeratotic cutaneous diseases
CC   including winter xerosis, psoriasis and various ichthyoses. Accordingly
CC   the dermatological activity of CDSN provides a good diagnostic and
CC   therapeutic target for alopecia, and through gene therapy CDSN can be
CC   used to modulate hair growth and treat male pattern baldness. This
CC   polypeptide sequence is the human corneodesmosin (CDSN) gene of the
CC   invention
XX
SQ   Sequence 529 AA;

  Query Match            100.0%;  Score 149;  DB 7;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 7
ADN04226
ID   ADN04226 standard; protein; 529 AA.
XX
AC   ADN04226;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #308.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04225.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 620; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 529 AA;

  Query Match            100.0%;  Score 149;  DB 8;  Length 529;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 8
ADP12525
ID   ADP12525 standard; protein; 529 AA.
XX
AC   ADP12525;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #135.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
```

```
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J,  Fry K,  Woodward R,  Ly N,  Prentice J,  Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 65; SEQ ID NO 2534; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   protein that is encoded by the mRNA of the invention.
XX
SQ   Sequence 529 AA;

  Query Match             100.0%;  Score 149;  DB 8;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-13;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 9
AAG64152
ID   AAG64152 standard; protein; 529 AA.
XX
AC   AAG64152;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Polypeptide #1 for vulgar psoriasis diagnosis.
XX
KW   MHC S; major histocompatibility complex S; vulgar psoriasis; diagnosis;
KW   primer; SEEK1; HCR; a-helix coiled-coil rod homologue; polymorphism.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 143
FT                   /label= Ser, Asn
FT   Misc-difference 151
FT                   /label= Ser, Cys
FT   Misc-difference 152
FT                   /label= Ser, Arg
FT   Misc-difference 200
FT                   /label= Ser, Phe
FT   Misc-difference 282
FT                   /label= Pro, Thr
FT   Misc-difference 401
FT                   /label= Ser, Gly
FT   Misc-difference 408
FT                   /label= Ala, Ser
FT   Misc-difference 409
FT                   /label= Gly, Val
FT   Misc-difference 410
FT                   /label= Ser, Leu
FT   Misc-difference 453
FT                   /label= Ser, Asn
FT   Misc-difference 527
FT                   /label= Asp, Asn
XX
PN   WO200142458-A1.
XX
PD   14-JUN-2001.
XX
PF   06-DEC-2000; 2000WO-JP008624.
XX
PR   06-DEC-1999;   99JP-00346867.
XX
PA   (INOK/) INOKO H.
XX
PI   Inoko H,  Tamiya G;
XX
DR   WPI; 2001-381680/40.
DR   N-PSDB; AAH73624.
XX
PT   New primer DNA, useful for detecting vulgar psoriasis.
XX
PS   Disclosure; Fig 2; 106pp; Japanese.
XX
CC   The invention relates to a method of diagnosing vulgar psoriasis using
CC   primers based on the sequences of the human MHC S, SEEK1 and HCR genes.
CC   By analysing the sequences of these genes in Japanese patients with
CC   psoriasis and in normal subjects, it has been found that some of the
CC   examined polymorphisms correlate significantly to the group of patients
CC   with psoriasis. Vulgar psoriasis can therefore be diagnosed by analysing
CC   these gene polymorphisms. The present sequence is provided in the
CC   specification
XX
SQ   Sequence 529 AA;

  Query Match             95.3%;  Score 142;  DB 4;  Length 529;
  Best Local Similarity   96.6%;  Pred. No. 2.8e-12;
  Matches   28;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||| ||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLXSP 529


RESULT 10
ADW17250
ID   ADW17250 standard; protein; 704 AA.
XX
AC   ADW17250;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Eucalyptus grandis transcription factor protein C3H-type family Seq 1000.
XX
KW   plant; transcription; gene regulation; gene expression; transgenic plant;
KW   drought resistance; disease resistance; salt tolerance; cold tolerance;
KW   freezing tolerance; flowering; flavor enhancer; flower color.
XX
OS   Eucalyptus grandis.
XX
PN   WO2005001050-A2.
```

XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-US017965.
XX
PR 06-JUN-2003; 2003US-0476189P.
XX
PA (ARBO-) ARBORGEN LLC.
XX
PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
DR WPI; 2005-075542/08.
DR N-PSDB; ADW16464.
XX
PT New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 31; SEQ ID NO 1000; 1265pp; English.
XX
CC This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/ frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polypeptide is a plant
CC transcription factor protein sequence of the invention.
XX
SQ Sequence 704 AA;

Query Match 41.3%; Score 61.5; DB 9; Length 704;
Best Local Similarity 54.5%; Pred. No. 4;
Matches 12; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

```
Qy          8 PLGPQLADPEVFLPQGELLDSP 29
              ||||  ::|::| || :|| ||
Db        165 PLGPN-SNPDMFCPQQDLLSSP 185
```

RESULT 11
ABB06906
ID ABB06906 standard; protein; 252 AA.
XX
AC ABB06906;
XX
DT 18-JUN-2002 (first entry)
XX
DE Micromonospora carbonacea everninomicin locus protein ORF 26.
XX
KW Micromonospora carbonacea; antibiotic; everninomicin; biosynthesis;
KW gene cluster; genetic manipulation; contig.
XX
OS Micromonospora carbonacea.
XX
PN WO200155180-A2.
XX
PD 02-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-CA000128.
XX
PR 27-JAN-2000; 2000US-0177711P.
XX
PA (ECOP-) ECOPIA BIOSCIENCES INC.
PA (FARN/) FARNET C.
XX
PI Staffa A, Zazopoulos E, Mercure S, Nowacki P;
XX
DR WPI; 2001-476185/51.
DR N-PSDB; ABL50559.
XX
PT Novel isolated gene cluster encoding polypeptides involved in
PT everninomicin biosynthesis useful for construction of everninomicin
PT overproducing strains, and to allow chemical modifications of
PT everninomicin to enhance certain properties.
XX
PS Claim 15; Fig 1; 181pp; English.
XX
CC ABL50555 to ABL50562 represent contigs 1 to 8 from the Micromonospora
CC carbonacea everninomicin biosynthetic locus gene cluster. The contigs
CC encode the protein sequences designated ORF (open reading frame) 1 to 49,
CC given in ABB06881 to ABB06930. The gene cluster is useful for the
CC construction of the everninomicin antibiotic in overproducing strains,
CC and to allow chemical modifications of everninomicin to enhance certain
CC properties via genetic manipulation or combinational biosynthesis. The
CC gene cluster can be used to produce genetic systems and genes encoding
CC novel enzyme activities, and avoid the problems of low yield and quality
CC of everninomicins produced by chemical synthesis
XX
SQ Sequence 252 AA;

Query Match 39.6%; Score 59; DB 4; Length 252;
Best Local Similarity 59.1%; Pred. No. 3;
Matches 13; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

```
Qy          8 PLGPQLADPEVFLPQGELLDSP 29
              || ||:  || ::|  ||||||
Db         70 PLLPQIIKPEKYVP--ELLDSP 89
```

RESULT 12
ABP99287
ID ABP99287 standard; protein; 252 AA.
XX
AC ABP99287;
XX
DT 23-OCT-2003 (revised)
DT 21-MAR-2003 (first entry)
XX
DE Orthosomycin biosynthetic polypeptide SEQ ID NO 161.
XX
KW Orthosomycin; biosynthesis; everninomicin; avilamycin; enzyme.
XX
OS Micromonospora carbonacea; aurantiaca.
XX
PN WO200279505-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-CA000432.
XX
PR 28-MAR-2001; 2001US-0279095P.
PR 30-MAR-2001; 2001US-0279709P.
PR 20-APR-2001; 2001US-0285214P.
XX
PA (ECOP-) ECOPIA BIOSCIENCES INC.
XX
PI Farnet CM, Zazopoulos E, Staffa A;
XX
DR WPI; 2003-058435/05.
DR N-PSDB; ABZ66750.
XX
PT Identifying orthosomycin biosynthetic gene, gene fragment or gene
PT cluster, by detecting presence of nucleic acid sequence corresponding to
PT 17 of flambamycins protein families.

```
XX
PS   Claim 1; Page 309-310; 511pp; English.
XX
CC   The invention relates to identifying orthosomycin biosynthetic genes and
CC   its fragment/gene cluster (ABZ66670-ABZ666813), comprising detecting the
CC   presence of a nucleic acid sequence coding for a polypeptide (ABP99207-
CC   ABP99362). The method is useful for identifying an orthosomycin
CC   biosynthetic gene, gene fragment or gene cluster, especially an
CC   everninomicin-type or avilamycin-type orthosomycin biosynthetic gene,
CC   gene fragment or gene cluster. The method is useful for detecting the
CC   presence of any organism that contains DNA for the production of
CC   orthosomycins (both everninomicin-type orthosomycins and avilamycin-type
CC   orthosomycins) regardless of the level at which genes for orthosomycin
CC   production are expressed by the organism or the amount of orthosomycin
CC   produced by the organism. This allows for the detection of new
CC   orthosomycin natural products, not produced by the organism. (Updated on
CC   23-OCT-2003 to standardise OS field)
XX
SQ   Sequence 252 AA;

  Query Match             39.6%;  Score 59;  DB 6;  Length 252;
  Best Local Similarity   59.1%;  Pred. No. 3;
  Matches   13;  Conservative    3;  Mismatches    4;  Indels    2;  Gaps    1;

Qy          8 PLGPQLADPEVFLPQGELLDSP 29
              || ||:  || ::|  ||||||
Db         70 PLLPQIIKPEKYVP--ELLDSP 89


RESULT 13
ABU36775
ID   ABU36775 standard; protein; 300 AA.
XX
AC   ABU36775;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #22302.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA40645.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 64699; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 300 AA;

  Query Match             37.6%;  Score 56;  DB 6;  Length 300;
  Best Local Similarity   41.4%;  Pred. No. 10;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        168 DVLAAALPAGP------AYYPDGELTDEP 190


RESULT 14
ABU34700
ID   ABU34700 standard; protein; 309 AA.
XX
AC   ABU34700;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #20227.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Mycobacterium bovis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA38570.
XX
```

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 62624; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 309 AA;

Query Match 37.6%; Score 56; DB 6; Length 309;
Best Local Similarity 41.4%; Pred. No. 11;
Matches 12; Conservative 2; Mismatches 9; Indels 6; Gaps 1;

```
Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        177 DVLAAALPAGP------AYYPDGELTDEP 199
```

RESULT 15
AAY83764
ID AAY83764 standard; protein; 272 AA.
XX
AC AAY83764;
XX
DT 21-JUN-2000 (first entry)
XX
DE Encephalitozoon hellem polar tube protein PTP35.
XX
KW Vaccine; gene therapy; polar tube protein; microsporidium; antibody;
KW antiparasitic; antiprotozoan; immunostimmulatory; immunoassay;
KW opportunistic pathogen; immunocompromised patient; AIDS.
XX
OS Encephalitozoon hellem.
XX
PN WO200001724-A1.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99WO-FR001630.
XX
PR 07-JUL-1998; 98FR-00008692.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Delbac F, Vivares C, Danchin A;
XX
DR WPI; 2000-237186/20.
DR N-PSDB; AAZ91196.
XX
PT New microsporidial tube protein and related nucleic acid, useful e.g. in
PT vaccines against Encephalitozoon infection and for raising diagnostic
PT antibodies.
XX
PS Claim 11; Page 46-47; 55pp; French.
XX
CC This sequence represents the complete polar tube protein PTP35 from the
CC microsporidium Encephalitozoon hellem. The protein has antiparasitic,
CC antiprotozoal and immunostimmulatory activities. The proteins are used to
CC raise specific antibodies, useful as immunoassay reagents for detecting
CC serum antibodies against Encephalitozoon (opportunistic pathogens in
CC immunocompromised, especially acquired immune deficiency syndrome,
CC patients), in vaccines to prevent such infections
XX
SQ Sequence 272 AA;

Query Match 36.9%; Score 55; DB 3; Length 272;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 15; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

```
Qy          4 AQVKPLGPQLADPEVFLP----------------------QGELLDSP 29
              ||| |: || | |  |||                        ||:: | |
Db         14 AQVAPVTPQAAVPTQFLPGAQQKIGGVDNRCANKQVEGVQIFQGDMADCP 63
```

Search completed: March 21, 2006, 21:45:57
Job time : 95.3273 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:43 ; Search time 75.9273 Seconds
(without alignments)
159.588 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	100.0	58	4	US-10-316-175-70	Sequence 70, Appl
2	149	100.0	529	4	US-10-316-175-15	Sequence 15, Appl
3	149	100.0	529	5	US-10-883-805-6	Sequence 6, Appli
4	149	100.0	529	5	US-10-677-865-1	Sequence 1, Appli
5	142	95.3	529	4	US-10-164-230-61	Sequence 61, Appl
6	142	95.3	529	4	US-10-204-884-3	Sequence 3, Appli
7	59	39.6	252	3	US-09-769-734-31	Sequence 31, Appl
8	59	39.6	252	4	US-10-107-431-161	Sequence 161, App
9	56	37.6	300	4	US-10-282-122A-64699	Sequence 64699, A
10	56	37.6	309	4	US-10-282-122A-62624	Sequence 62624, A
11	55	36.9	272	3	US-09-755-456-10	Sequence 10, Appl
12	54.5	36.6	410	5	US-10-501-282-6636	Sequence 6636, Ap
13	54.5	36.6	428	5	US-10-501-282-6638	Sequence 6638, Ap
14	54.5	36.6	433	5	US-10-501-282-6640	Sequence 6640, Ap
15	54	36.2	299	4	US-10-282-122A-62182	Sequence 62182, A
16	53	35.6	84	4	US-10-425-115-354462	Sequence 354462,
17	53	35.6	129	4	US-10-425-115-368681	Sequence 368681,
18	52	34.9	312	5	US-10-450-763-34434	Sequence 34434, A
19	51.5	34.6	837	4	US-10-369-493-22296	Sequence 22296, A
20	51.5	34.6	837	5	US-10-732-923-23246	Sequence 23246, A
21	51	34.2	108	4	US-10-425-115-188845	Sequence 188845,
22	51	34.2	229	4	US-10-425-115-226842	Sequence 226842,
23	51	34.2	4561	5	US-10-450-763-60562	Sequence 60562, A
24	51	34.2	9222	5	US-10-450-763-51423	Sequence 51423, A
25	50	33.6	78	4	US-10-425-115-353935	Sequence 353935,
26	50	33.6	84	4	US-10-437-963-125244	Sequence 125244,
27	50	33.6	248	4	US-10-425-115-337597	Sequence 337597,
28	50	33.6	275	4	US-10-425-115-337608	Sequence 337608,
29	50	33.6	283	4	US-10-425-114-65239	Sequence 65239, A
30	50	33.6	289	4	US-10-314-657-2	Sequence 2, Appli
31	50	33.6	289	5	US-10-473-193-2	Sequence 2, Appli
32	50	33.6	353	5	US-10-450-763-37710	Sequence 37710, A
33	50	33.6	355	5	US-10-450-763-49307	Sequence 49307, A
34	50	33.6	400	4	US-10-369-493-20647	Sequence 20647, A
35	49.5	33.2	70	4	US-10-425-115-364294	Sequence 364294,
36	49	32.9	46	4	US-10-600-816-90	Sequence 90, Appl
37	49	32.9	117	4	US-10-425-115-297868	Sequence 297868,
38	49	32.9	167	4	US-10-425-115-312427	Sequence 312427,
39	49	32.9	345	4	US-10-225-567A-619	Sequence 619, App
40	49	32.9	345	4	US-10-467-252-8	Sequence 8, Appli
41	49	32.9	345	4	US-10-600-816-4	Sequence 4, Appli
42	49	32.9	345	5	US-10-500-428-2	Sequence 2, Appli
43	49	32.9	347	4	US-10-437-963-130021	Sequence 130021,
44	49	32.9	362	4	US-10-182-822A-14	Sequence 14, Appl
45	49	32.9	448	4	US-10-369-493-21148	Sequence 21148, A

ALIGNMENTS

RESULT 1
US-10-316-175-70
; Sequence 70, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-316-175-70

Query Match 100.0%; Score 149; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db         30 DILAQVKPLGPQLADPEVFLPQGELLDSP 58
```

RESULT 2
US-10-316-175-15
; Sequence 15, Application US/10316175
; Publication No. US20030211065A1
; GENERAL INFORMATION:
; APPLICANT: WELLA AG
; APPLICANT: Pras, Elon
; APPLICANT: Levy-Nissenbaum, Etgar
; APPLICANT: Frydman, Moshe
; APPLICANT: Thederahn, Ted
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING HYPOTRICHOSIS SIMPLEX
; FILE REFERENCE: 532862000100
; CURRENT APPLICATION NUMBER: US/10/316,175
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/338,188

```
;  PRIOR FILING DATE: 2001-12-07
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-316-175-15

  Query Match             100.0%;  Score 149;  DB 4;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 3
US-10-883-805-6
; Sequence 6, Application US/10883805
; Publication No. US20050048612A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  CORLEY, Neil C.
;  APPLICANT:  GUEGLER, Karl J.
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YUE, Henry
;  TITLE OF INVENTION: HUMAN EPIDERMAL PROTEINS
;  FILE REFERENCE: PF-0567 PCT
;  CURRENT APPLICATION NUMBER: US/10/883,805
;  CURRENT FILING DATE:  2004-07-06
;  PRIOR APPLICATION NUMBER: 09/123,531
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-07-28
;  PRIOR APPLICATION NUMBER: 09/206,817
;  PRIOR FILING DATE: 1998-12-07
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No: 3325323
US-10-883-805-6

  Query Match             100.0%;  Score 149;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 4
US-10-677-865-1
; Sequence 1, Application US/10677865
; Publication No. US20050112150A1
; GENERAL INFORMATION:
;  APPLICANT: SERRE, Guy Bruno Rene
;  APPLICANT:  SIMON, Michel
;  APPLICANT:  WEBER-VIVAT, Marina
;  TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
;  TITLE OF INVENTION:  AND USE THEREOF
;  FILE REFERENCE: 016800-336
;  CURRENT APPLICATION NUMBER: US/10/677,865
;  CURRENT FILING DATE:  2003-10-03
;  PRIOR APPLICATION NUMBER: US/09/381,656
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: FR 97/03899
;  PRIOR FILING DATE: 1997-03-28
;  PRIOR APPLICATION NUMBER: FR 97/11317
;  PRIOR FILING DATE: 1997-09-11
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Human
US-10-677-865-1

  Query Match             100.0%;  Score 149;  DB 5;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-12;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |||||||||||||||||||||||||||||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529


RESULT 5
US-10-164-230-61
; Sequence 61, Application US/10164230
; Publication No. US20030170652A1
; GENERAL INFORMATION:
;  APPLICANT: Inoko, Hidetoshi
;  APPLICANT:  Tamiya, Gen
;  TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
;  FILE REFERENCE: 06501-112US1
;  CURRENT APPLICATION NUMBER: US/10/164,230
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: PCT/JP00/08624
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: JP 11/346867
;  PRIOR FILING DATE: 1999-12-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
```

; LOCATION: 408
; OTHER INFORMATION: Xaa = Ala or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 409
; OTHER INFORMATION: Xaa = Gly or Val
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 410
; OTHER INFORMATION: Xaa = Ser or Leu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 453
; OTHER INFORMATION: Xaa = Ser or Asn
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 527
; OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

Query Match 95.3%; Score 142; DB 4; Length 529;
Best Local Similarity 96.6%; Pred. No. 3.3e-11;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
||||||||||||||||||||||||||| ||
Db 501 DILAQVKPLGPQLADPEVFLPQGELLXSP 529

RESULT 6
US-10-204-884-3
; Sequence 3, Application US/10204884
; Publication No. US20030186371A1
; GENERAL INFORMATION:
; APPLICANT: Oxagen Limited
; APPLICANT: Olaveson, Mark
; APPLICANT: Lench, Nick
; APPLICANT: Allen, Maxine
; APPLICANT: Tazi-Ahnini, Rachind
; TITLE OF INVENTION: Test and model for inflammatory disease
; FILE REFERENCE: P30000WO-PS
; CURRENT APPLICATION NUMBER: US/10/204,884
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: GB 0004312.5
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (18)..(18)
; OTHER INFORMATION: The 'Xaa' at location 18 stands for Met, or Leu.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (56)..(56)
; OTHER INFORMATION: The 'Xaa' at location 56 stands for Leu, or Phe.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (143)..(143)
; OTHER INFORMATION: The 'Xaa' at location 143 stands for Ser, or Asn.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (202)..(202)
; OTHER INFORMATION: The 'Xaa' at location 202 stands for Ser, or Phe.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (401)..(401)
; OTHER INFORMATION: The 'Xaa' at location 401 stands for Gly, or Ser.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (408)..(408)
; OTHER INFORMATION: The 'Xaa' at location 408 stands for Ala, or Ser.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (409)..(409)
; OTHER INFORMATION: The 'Xaa' at location 409 stands for Gly, or Val.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (410)..(410)
; OTHER INFORMATION: The 'Xaa' at location 410 stands for Ser, or Leu.
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (527)..(527)
; OTHER INFORMATION: The 'Xaa' at location 527 stands for Asp, or Asn.
US-10-204-884-3

Query Match 95.3%; Score 142; DB 4; Length 529;
Best Local Similarity 96.6%; Pred. No. 3.3e-11;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
||||||||||||||||||||||||||| ||
Db 501 DILAQVKPLGPQLADPEVFLPQGELLXSP 529

RESULT 7
US-09-769-734-31
; Sequence 31, Application US/09769734
; Publication No. US20030143666A1
; GENERAL INFORMATION:
; APPLICANT: Ecopia BioSciences Inc.
; TITLE OF INVENTION: Genetic Locus for Everninomicin Biosynthesis
; FILE REFERENCE: PA 005-US
; CURRENT APPLICATION NUMBER: US/09/769,734
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 252
; TYPE: PRT
; ORGANISM: M/ carbonacea
US-09-769-734-31

Query Match 39.6%; Score 59; DB 3; Length 252;
Best Local Similarity 59.1%; Pred. No. 5.9;
Matches 13; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

Qy 8 PLGPQLADPEVFLPQGELLDSP 29
|| ||: || ::| ||||||
Db 70 PLLPQIIKPEKYVP--ELLDSP 89

RESULT 8
US-10-107-431-161
; Sequence 161, Application US/10107431
; Publication No. US20030224364A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AND DISTINGUISHING ORTH
; TITLE OF INVENTION: BIOSYNTHETIC LOCI
; FILE REFERENCE: 3001-7US
; CURRENT APPLICATION NUMBER: US/10/107,431
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 161
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Micromonospora carbonacea aurantiaca
US-10-107-431-161

```
  Query Match            39.6%;  Score 59;  DB 4;  Length 252;
  Best Local Similarity  59.1%;  Pred. No. 5.9;
  Matches   13;  Conservative    3;  Mismatches    4;  Indels    2;  Gaps    1;

Qy          8 PLGPQLADPEVFLPQGELLDSP 29
              || ||:  || ::|  ||||||
Db         70 PLLPQIIKPEKYVP--ELLDSP 89


RESULT 9
US-10-282-122A-64699
; Sequence 64699, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64699
;   LENGTH: 300
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64699

  Query Match            37.6%;  Score 56;  DB 4;  Length 300;
  Best Local Similarity  41.4%;  Pred. No. 19;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        168 DVLAAALPAGP------AYYPDGELTDEP 190


RESULT 10
US-10-282-122A-62624
; Sequence 62624, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62624
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Mycobacterium bovis
US-10-282-122A-62624

  Query Match            37.6%;  Score 56;  DB 4;  Length 309;
  Best Local Similarity  41.4%;  Pred. No. 19;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        177 DVLAAALPAGP------AYYPDGELTDEP 199


RESULT 11
US-09-755-456-10
; Sequence 10, Application US/09755456
; Patent No. US20010021512A1
; GENERAL INFORMATION:
;  APPLICANT: DELBAC, FREDERIC
;  APPLICANT:  DANCHIN, ANTOINE
;  APPLICANT:  VIVARES, CHRISTIAN
;  TITLE OF INVENTION: MICROSPORIDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS
;  TITLE OF INVENTION:  CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
;  FILE REFERENCE: 1566-00
;  CURRENT APPLICATION NUMBER: US/09/755,456
;  CURRENT FILING DATE:  2001-01-05
;  PRIOR APPLICATION NUMBER: PCT/FR99/01630
;  PRIOR FILING DATE: 1999-07-06
;  PRIOR APPLICATION NUMBER: FR 98/08692
;  PRIOR FILING DATE: 1998-07-07
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 10
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Encephalitozoon hellem
US-09-755-456-10

  Query Match             36.9%;  Score 55;  DB 3;  Length 272;
  Best Local Similarity   30.0%;  Pred. No. 23;
  Matches   15;  Conservative    3;  Mismatches    8;  Indels   24;  Gaps    1;

Qy          4 AQVKPLGPQLADPEVFLP-----------------------QGELLDSP 29
              ||| |: || | |  |||                       ||:: | |
Db         14 AQVAPVTPQAAVPTQFLPGAQQKIGGVDNRCANKQVEGVQIFQGDMADCP 63


RESULT 12
US-10-501-282-6636
; Sequence 6636, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
;  APPLICANT: MCMICHAEL, JOHN CALHOUN
;  APPLICANT:  ZAGURSKY, ROBERT JOHN
;  APPLICANT:  RUSSELL, DAVID PARRISH
;  APPLICANT:  FLETCHER, LEAH DIANE
;  TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
;  TITLE OF INVENTION:  POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
;  FILE REFERENCE: AM100780 L2
;  CURRENT APPLICATION NUMBER: US/10/501,282
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: 60/333,777
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: 60/426,742
;  PRIOR FILING DATE: 2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US02/36123
;  PRIOR FILING DATE: 2002-11-25
;  NUMBER OF SEQ ID NOS: 6653
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6636
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Alloiococcus otitidis
US-10-501-282-6636

  Query Match             36.6%;  Score 54.5;  DB 5;  Length 410;
  Best Local Similarity   50.0%;  Pred. No. 43;
  Matches   15;  Conservative    3;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFL-PQGELLDSP 29
              ||    :|   :| |||  | || ||:|||
Db        353 DIPEDSRPALERLLDPEEALAPQYELVDSP 382


RESULT 13
US-10-501-282-6638
; Sequence 6638, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
;  APPLICANT: MCMICHAEL, JOHN CALHOUN
;  APPLICANT:  ZAGURSKY, ROBERT JOHN
;  APPLICANT:  RUSSELL, DAVID PARRISH
;  APPLICANT:  FLETCHER, LEAH DIANE
;  TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
;  TITLE OF INVENTION:  POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
;  FILE REFERENCE: AM100780 L2
;  CURRENT APPLICATION NUMBER: US/10/501,282
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: 60/333,777
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: 60/426,742
;  PRIOR FILING DATE: 2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US02/36123
;  PRIOR FILING DATE: 2002-11-25
;  NUMBER OF SEQ ID NOS: 6653
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6638
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Alloiococcus otitidis
US-10-501-282-6638

  Query Match             36.6%;  Score 54.5;  DB 5;  Length 428;
  Best Local Similarity   50.0%;  Pred. No. 45;
  Matches   15;  Conservative    3;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFL-PQGELLDSP 29
              ||    :|   :| |||  | || ||:|||
Db        371 DIPEDSRPALERLLDPEEALAPQYELVDSP 400


RESULT 14
US-10-501-282-6640
; Sequence 6640, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
;  APPLICANT: MCMICHAEL, JOHN CALHOUN
;  APPLICANT:  ZAGURSKY, ROBERT JOHN
;  APPLICANT:  RUSSELL, DAVID PARRISH
;  APPLICANT:  FLETCHER, LEAH DIANE
;  TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
;  TITLE OF INVENTION:  POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
;  FILE REFERENCE: AM100780 L2
;  CURRENT APPLICATION NUMBER: US/10/501,282
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: 60/333,777
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: 60/426,742
;  PRIOR FILING DATE: 2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US02/36123
;  PRIOR FILING DATE: 2002-11-25
;  NUMBER OF SEQ ID NOS: 6653
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6640
;   LENGTH: 433
;   TYPE: PRT
;   ORGANISM: Alloiococcus otitidis
US-10-501-282-6640

  Query Match             36.6%;  Score 54.5;  DB 5;  Length 433;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches   15;  Conservative    3;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFL-PQGELLDSP 29
              ||    :|   :| |||  | || ||:|||
Db        376 DIPEDSRPALERLLDPEEALAPQYELVDSP 405


RESULT 15
US-10-282-122A-62182
; Sequence 62182, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
```

```
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62182
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: Mycobacterium avium
US-10-282-122A-62182

  Query Match             36.2%;  Score 54;  DB 4;  Length 299;
  Best Local Similarity   41.4%;  Pred. No. 36;
  Matches   12;  Conservative    2;  Mismatches    9;  Indels    6;  Gaps    1;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              |:||   | ||       : | ||| | |
Db        167 DVLAAALPPGP------AYYPDGELTDEP 189


Search completed: March 21, 2006, 21:52:53
Job time : 75.9273 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:50:33 ; Search time 12.6545 Seconds
(without alignments)
65.594 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	32.2	177	7	US-11-131-744-3	Sequence 3, Appli
2	48	32.2	356	7	US-11-096-568A-22843	Sequence 22843, A
3	47	31.5	545	6	US-10-714-995-8	Sequence 8, Appli
4	46	30.9	258	7	US-11-000-463-871	Sequence 871, App
5	46	30.9	437	7	US-11-087-099-6889	Sequence 6889, Ap
6	46	30.9	511	7	US-11-098-686-10931	Sequence 10931, A
7	45.5	30.5	338	7	US-11-098-686-10126	Sequence 10126, A
8	45.5	30.5	517	7	US-11-096-568A-13312	Sequence 13312, A
9	45.5	30.5	527	7	US-11-096-568A-13311	Sequence 13311, A
10	45	30.2	330	7	US-11-072-512-2839	Sequence 2839, Ap
11	45	30.2	386	7	US-11-087-099-7489	Sequence 7489, Ap
12	44	29.5	343	6	US-10-497-135-11	Sequence 11, Appl
13	44	29.5	343	6	US-10-497-135-12	Sequence 12, Appl
14	44	29.5	387	6	US-10-485-517-132	Sequence 132, App
15	44	29.5	455	7	US-11-087-099-404	Sequence 404, App
16	44	29.5	458	6	US-10-511-989-166	Sequence 166, App
17	44	29.5	701	6	US-10-517-939-6	Sequence 6, Appli
18	44	29.5	810	6	US-10-510-716-9	Sequence 9, Appli
19	43.5	29.2	332	6	US-10-793-626-2762	Sequence 2762, Ap
20	43.5	29.2	391	7	US-11-072-512-3849	Sequence 3849, Ap
21	43.5	29.2	399	7	US-11-072-512-2799	Sequence 2799, Ap
22	43.5	29.2	879	7	US-11-098-686-10290	Sequence 10290, A
23	43.5	29.2	3475	7	US-11-087-099-10885	Sequence 10885, A
24	43	28.9	43	7	US-11-131-744-4	Sequence 4, Appli
25	43	28.9	45	6	US-10-957-887B-72	Sequence 72, Appl
26	43	28.9	142	6	US-10-986-501-270	Sequence 270, App
27	43	28.9	444	7	US-11-087-099-956	Sequence 956, App
28	43	28.9	1308	6	US-10-912-971-12	Sequence 12, Appl
29	43	28.9	1308	7	US-11-113-202-16	Sequence 16, Appl
30	43	28.9	2556	7	US-11-050-346-67	Sequence 67, Appl
31	42.5	28.5	69	5	US-09-978-360A-704	Sequence 704, App
32	42.5	28.5	81	6	US-10-821-234-1613	Sequence 1613, Ap
33	42.5	28.5	716	6	US-10-330-773-355	Sequence 355, App
34	42.5	28.5	742	7	US-11-087-099-7459	Sequence 7459, Ap
35	42.5	28.5	790	6	US-10-330-773-358	Sequence 358, App
36	42.5	28.5	1212	7	US-11-080-991-74	Sequence 74, Appl
37	42.5	28.5	1212	7	US-11-101-000-1	Sequence 1, Appli
38	42.5	28.5	1212	7	US-11-186-284-191	Sequence 191, App
39	42.5	28.5	1212	7	US-11-130-945-1	Sequence 1, Appli
40	42	28.2	148	7	US-11-096-568A-10263	Sequence 10263, A
41	42	28.2	162	7	US-11-096-568A-10262	Sequence 10262, A
42	42	28.2	164	7	US-11-087-099-6508	Sequence 6508, Ap
43	42	28.2	165	7	US-11-087-099-9748	Sequence 9748, Ap
44	42	28.2	203	7	US-11-096-568A-10261	Sequence 10261, A
45	42	28.2	326	7	US-11-096-568A-19445	Sequence 19445, A

ALIGNMENTS

RESULT 1
US-11-131-744-3
; Sequence 3, Application US/11131744
; Publication No. US20060023808A1
; GENERAL INFORMATION:
; APPLICANT: HAJIVANDI, MAHBOD R.
; APPLICANT: POPE, ROBERT M.
; APPLICANT: SADEGHI, MEHRNOOSH
; APPLICANT: LIANG, XIQUAN
; APPLICANT: LEITE, JOHN F.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR CALIBRATION IN MASS
; TITLE OF INVENTION: SPECTROMETRY
; FILE REFERENCE: NVX-28 US
; CURRENT APPLICATION NUMBER: US/11/131,744
; CURRENT FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: 60/572,290
; PRIOR FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: 60/669,373
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 60/572,215
; PRIOR FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: 60/621,685
; PRIOR FILING DATE: 2004-10-26
; PRIOR APPLICATION NUMBER: 60/621,686
; PRIOR FILING DATE: 2004-10-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 3
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: protein sequence
US-11-131-744-3

Query Match 32.2%; Score 48; DB 7; Length 177;
Best Local Similarity 50.0%; Pred. No. 8.5;
Matches 12; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```
Qy          3 LAQVKPLGPQLADPEVFLPQGELL 26
              | | : |  :| :||| ||  |||
Db        113 LDQYRALLGKLTNPEVELPNAELL 136
```

RESULT 2
US-11-096-568A-22843
; Sequence 22843, Application US/11096568A

```
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22843
;   LENGTH: 356
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(356)
;   OTHER INFORMATION: Ceres Seq. ID no. 12410333
US-11-096-568A-22843

  Query Match             32.2%;  Score 48;  DB 7;  Length 356;
  Best Local Similarity   44.4%;  Pred. No. 19;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 QLADPEVFLPQGELLDSP 29
              :: ||:  | ||::|| |
Db        241 KITDPKTSLQQGQILDVP 258


RESULT 3
US-10-714-995-8
; Sequence 8, Application US/10714995
; Publication No. US20060039941A1
; GENERAL INFORMATION:
;  APPLICANT: Stam, Lynn
;  APPLICANT:  Kamdar, Kim
;  APPLICANT:  Spana, Eric
;  APPLICANT:  Bachman, Jane
;  TITLE OF INVENTION: Nucleic Acid Sequences from Drosophila melanogaster that encode P
;  TITLE OF INVENTION:  Essential for Viability and Uses Thereof
;  FILE REFERENCE: 70201USNP
;  CURRENT APPLICATION NUMBER: US/10/714,995
;  CURRENT FILING DATE:  2003-11-17
;  PRIOR APPLICATION NUMBER: US 60/436,442
;  PRIOR FILING DATE: 2002-12-23
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 545
;   TYPE: PRT
;   ORGANISM: drosophila melanogaster
US-10-714-995-8

  Query Match             31.5%;  Score 47;  DB 6;  Length 545;
  Best Local Similarity   47.4%;  Pred. No. 43;
  Matches    9;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         11 PQLADPEVFLPQGELLDSP 29
              ||| |  :  :| | | | |
Db        257 PQLPDTDAVMPNGWLEDEP 275


RESULT 4
US-11-000-463-871
; Sequence 871, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 871
;   LENGTH: 258
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-000-463-871

  Query Match             30.9%;  Score 46;  DB 7;  Length 258;
  Best Local Similarity   35.7%;  Pred. No. 25;
  Matches   10;  Conservative    6;  Mismatches   10;  Indels    2;  Gaps    1;

Qy          3 LAQVKPLGPQLADP--EVFLPQGELLDS 28
              |  :||| | : |   ::   ||:: ||
Db         26 LTDIKPLPPNIKDRLIKIMSMQGQITDS 53


RESULT 5
US-11-087-099-6889
; Sequence 6889, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6889
;   LENGTH: 437
;   TYPE: PRT
;   ORGANISM: Coxiella burnetii RSA 493
US-11-087-099-6889

  Query Match             30.9%;  Score 46;  DB 7;  Length 437;
  Best Local Similarity   34.8%;  Pred. No. 47;
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 AQVKPLGPQLADPEVFLPQGELL 26
              |::  || :: :|: |:||  :|
Db        219 ARIATLGEEVKNPKRFIPQAIIL 241


RESULT 6
US-11-098-686-10931
; Sequence 10931, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
;  APPLICANT: Kapur, Vivek and Gebhart, Connie J.
;  TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
```

```
;  FILE REFERENCE: 09531-128001
;  CURRENT APPLICATION NUMBER: US/11/098,686
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: PCT/US03/31318
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: US 60/416,395
;  PRIOR FILING DATE: 2002-10-04
;  NUMBER OF SEQ ID NOS: 11433
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10931
;   LENGTH: 511
;   TYPE: PRT
;   ORGANISM: Lawsonia intracellularis
US-11-098-686-10931

  Query Match             30.9%;  Score 46;  DB 7;  Length 511;
  Best Local Similarity   37.0%;  Pred. No. 56;
  Matches   10;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 ILAQVKPLGPQLADPEVFLPQGELLDS 28
              |: :  ||| :  || ||: :|   |:
Db        258 IILEYAPLGHENDDPLVFVGKGITFDT 284


RESULT 7
US-11-098-686-10126
; Sequence 10126, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
;  APPLICANT: Kapur, Vivek and Gebhart, Connie J.
;  TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
;  FILE REFERENCE: 09531-128001
;  CURRENT APPLICATION NUMBER: US/11/098,686
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: PCT/US03/31318
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: US 60/416,395
;  PRIOR FILING DATE: 2002-10-04
;  NUMBER OF SEQ ID NOS: 11433
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10126
;   LENGTH: 338
;   TYPE: PRT
;   ORGANISM: Lawsonia intracellularis
US-11-098-686-10126

  Query Match             30.5%;  Score 45.5;  DB 7;  Length 338;
  Best Local Similarity   46.9%;  Pred. No. 41;
  Matches   15;  Conservative    3;  Mismatches   11;  Indels    3;  Gaps    2;

Qy          1 DILAQVKPLGPQLADP--EVFLPQG-ELLDSP 29
              | |  :  || || :   |: || | |||| |
Db        229 DALKPLTELGIQLGESNMELNLPNGIELLDVP 260


RESULT 8
US-11-096-568A-13312
; Sequence 13312, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 13312
;   LENGTH: 517
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(517)
;   OTHER INFORMATION: Ceres Seq. ID no. 15174869
US-11-096-568A-13312

  Query Match             30.5%;  Score 45.5;  DB 7;  Length 517;
  Best Local Similarity   42.3%;  Pred. No. 67;
  Matches   11;  Conservative    5;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          3 LAQVKPLGPQLADPEVFLPQG-ELLD 27
              :|  ||  | : | :||  || |::|
Db         13 VADAKPQCPAVLDADVFRRQGHEVID 38


RESULT 9
US-11-096-568A-13311
; Sequence 13311, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 13311
;   LENGTH: 527
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(527)
;   OTHER INFORMATION: Ceres Seq. ID no. 15174868
US-11-096-568A-13311

  Query Match             30.5%;  Score 45.5;  DB 7;  Length 527;
  Best Local Similarity   42.3%;  Pred. No. 68;
  Matches   11;  Conservative    5;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          3 LAQVKPLGPQLADPEVFLPQG-ELLD 27
              :|  ||  | : | :||  || |::|
Db         23 VADAKPQCPAVLDADVFRRQGHEVID 48


RESULT 10
US-11-072-512-2839
; Sequence 2839, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
```

```
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2839
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2839

  Query Match             30.2%;  Score 45;  DB 7;  Length 330;
  Best Local Similarity   37.5%;  Pred. No. 47;
  Matches    9;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          6 VKPLGPQLADPEVFLPQGELLDSP 29
              ::| |  |  |:::||   ||  |
Db         86 LQPRGAHLPPPQLWLPPQALLGRP 109


RESULT 11
US-11-087-099-7489
; Sequence 7489, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7489
;   LENGTH: 386
;   TYPE: PRT
;   ORGANISM: Xanthomonas campestris pv. campestris str. ATCC 33913
US-11-087-099-7489

  Query Match             30.2%;  Score 45;  DB 7;  Length 386;
  Best Local Similarity   52.6%;  Pred. No. 56;
  Matches   10;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 GPQLADPEVFLPQGELLDS 28
              |    |||:|||   | ||
Db        273 GRHFPDPEIFLPARWLDDS 291


RESULT 12
US-10-497-135-11
; Sequence 11, Application US/10497135
; Publication No. US20050272132A1
; GENERAL INFORMATION:
;  APPLICANT: Biotica Technology Ltd
;  APPLICANT:  Gregory, Matthew A
;  APPLICANT:  Gaisser, Sabine
;  APPLICANT:  Petkovic, Hrvoje
;  APPLICANT:  Moss, Steven
;  TITLE OF INVENTION: Production of Polyketides and Other Natural Products
;  FILE REFERENCE: 4408-P03444US00
;  CURRENT APPLICATION NUMBER: US/10/497,135
;  CURRENT FILING DATE:  2004-05-26
;  PRIOR APPLICATION NUMBER: PCT/GB03/003230
;  PRIOR FILING DATE: 2003-07-16
;  PRIOR APPLICATION NUMBER: GB0216509.0
;  PRIOR FILING DATE: 2002-07-16
;  PRIOR APPLICATION NUMBER: GB0224922.5
;  PRIOR FILING DATE: 2002-10-25
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Streptomyces hygroscopicus
US-10-497-135-11

  Query Match             29.5%;  Score 44;  DB 6;  Length 343;
  Best Local Similarity   45.8%;  Pred. No. 68;
  Matches   11;  Conservative    2;  Mismatches    7;  Indels    4;  Gaps    1;

Qy          7 KPLGPQLAD----PEVFLPQGELL 26
              : ||| |||    ||:  |  | |
Db        273 RELGPSLADLCAAPEIAAPHPERL 296


RESULT 13
US-10-497-135-12
; Sequence 12, Application US/10497135
; Publication No. US20050272132A1
; GENERAL INFORMATION:
;  APPLICANT: Biotica Technology Ltd
;  APPLICANT:  Gregory, Matthew A
;  APPLICANT:  Gaisser, Sabine
;  APPLICANT:  Petkovic, Hrvoje
;  APPLICANT:  Moss, Steven
;  TITLE OF INVENTION: Production of Polyketides and Other Natural Products
;  FILE REFERENCE: 4408-P03444US00
;  CURRENT APPLICATION NUMBER: US/10/497,135
;  CURRENT FILING DATE:  2004-05-26
;  PRIOR APPLICATION NUMBER: PCT/GB03/003230
;  PRIOR FILING DATE: 2003-07-16
;  PRIOR APPLICATION NUMBER: GB0216509.0
;  PRIOR FILING DATE: 2002-07-16
;  PRIOR APPLICATION NUMBER: GB0224922.5
;  PRIOR FILING DATE: 2002-10-25
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Streptomyces hygroscopicus
US-10-497-135-12

  Query Match             29.5%;  Score 44;  DB 6;  Length 343;
  Best Local Similarity   45.8%;  Pred. No. 68;
  Matches   11;  Conservative    2;  Mismatches    7;  Indels    4;  Gaps    1;

Qy          7 KPLGPQLAD----PEVFLPQGELL 26
              : ||| |||    ||:  |  | |
Db        273 RELGPSLADLCAAPEIAAPHPERL 296


RESULT 14
US-10-485-517-132
; Sequence 132, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
;  APPLICANT: University of Sheffield
;  APPLICANT:  Biosynexus Incorporated
;  APPLICANT:  Foster, Simon
;  APPLICANT:  Mond, James
;  TITLE OF INVENTION: Antigenic Polypeptides
;  FILE REFERENCE: P100629WO
;  CURRENT APPLICATION NUMBER: US/10/485,517
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: GB 0118825.9
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: GB 0200349.9
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 424
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 132
;   LENGTH: 387
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-485-517-132

  Query Match             29.5%;  Score 44;  DB 6;  Length 387;
```

```
  Best Local Similarity   50.0%;  Pred. No. 78;
  Matches    8;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         11 PQLADPEVFLPQGELL 26
              ||:  |:| :|| :||
Db        243 PQIQSPKVEVPQSKLL 258


RESULT 15
US-11-087-099-404
; Sequence 404, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 404
;   LENGTH: 455
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-11-087-099-404

  Query Match             29.5%;  Score 44;  DB 7;  Length 455;
  Best Local Similarity   47.6%;  Pred. No. 95;
  Matches   10;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          9 LGPQLADPEVFLPQGELLDSP 29
              |||   | || | :|:  :||
Db        325 LGPLSEDQEVVLEEGKERNSP 345


Search completed: March 21, 2006, 21:53:23
Job time : 12.6545 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2006, 21:49:04 ; Search time 26.8909 Seconds
(without alignments)
89.160 Million cell updates/sec

Title: US-10-677-865A-5
Perfect score: 149
Sequence: 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	149	100.0	529	2	US-09-381-656-1	Sequence 1, Appli
2	142	95.3	529	2	US-10-164-230-61	Sequence 61, Appl
3	55	36.9	272	2	US-09-755-456-10	Sequence 10, Appl
4	52.5	35.2	272	2	US-09-949-016-8863	Sequence 8863, Ap
5	49	32.9	530	2	US-08-486-049-3	Sequence 3, Appli
6	49	32.9	530	2	US-10-314-739A-3	Sequence 3, Appli
7	48.5	32.6	337	2	US-08-686-528A-2	Sequence 2, Appli
8	48.5	32.6	337	2	US-09-456-287-2	Sequence 2, Appli
9	48	32.2	586	2	US-09-652-603-2	Sequence 2, Appli
10	48	32.2	692	2	US-09-758-759-79	Sequence 79, Appl
11	47.5	31.9	285	2	US-09-327-681-6	Sequence 6, Appli
12	47.5	31.9	425	2	US-09-252-991A-27848	Sequence 27848, A
13	47.5	31.9	563	2	US-09-431-470-4	Sequence 4, Appli
14	47	31.5	418	2	US-09-902-540-15577	Sequence 15577, A
15	47	31.5	427	2	US-09-252-991A-23349	Sequence 23349, A
16	46	30.9	398	2	US-09-902-540-12997	Sequence 12997, A
17	45.5	30.5	298	2	US-09-248-796A-16792	Sequence 16792, A
18	45.5	30.5	309	2	US-09-949-016-6485	Sequence 6485, Ap
19	45.5	30.5	318	2	US-09-949-016-7470	Sequence 7470, Ap
20	45.5	30.5	321	2	US-09-543-681A-4740	Sequence 4740, Ap
21	45.5	30.5	391	2	US-09-645-321-1	Sequence 1, Appli
22	45.5	30.5	391	2	US-10-378-745-1	Sequence 1, Appli
23	45.5	30.5	566	2	US-09-431-470-2	Sequence 2, Appli
24	45.5	30.5	1133	2	US-09-252-991A-17945	Sequence 17945, A
25	45	30.2	69	2	US-09-149-476-601	Sequence 601, App
26	45	30.2	70	2	US-09-149-476-465	Sequence 465, App
27	45	30.2	102	2	US-09-621-976-6071	Sequence 6071, Ap
28	45	30.2	113	2	US-09-513-999C-7675	Sequence 7675, Ap
29	45	30.2	147	2	US-09-489-039A-13311	Sequence 13311, A
30	45	30.2	157	2	US-09-248-796A-17506	Sequence 17506, A
31	45	30.2	159	2	US-09-252-991A-19997	Sequence 19997, A
32	45	30.2	174	2	US-08-833-167-48	Sequence 48, Appl
33	45	30.2	174	2	US-08-833-167-49	Sequence 49, Appl
34	45	30.2	174	2	US-08-833-167-50	Sequence 50, Appl
35	45	30.2	174	2	US-08-833-167-51	Sequence 51, Appl
36	45	30.2	174	2	US-08-833-167-52	Sequence 52, Appl
37	45	30.2	174	2	US-08-833-167-53	Sequence 53, Appl
38	45	30.2	174	2	US-08-833-167-54	Sequence 54, Appl
39	45	30.2	174	2	US-08-833-167-55	Sequence 55, Appl
40	45	30.2	174	2	US-08-833-167-56	Sequence 56, Appl
41	45	30.2	174	2	US-08-833-167-95	Sequence 95, Appl
42	45	30.2	174	2	US-08-833-167-96	Sequence 96, Appl
43	45	30.2	174	2	US-08-833-167-97	Sequence 97, Appl
44	45	30.2	174	2	US-08-833-167-98	Sequence 98, Appl
45	45	30.2	174	2	US-08-833-167-99	Sequence 99, Appl

ALIGNMENTS

RESULT 1
US-09-381-656-1
; Sequence 1, Application US/09381656
; Patent No. 6645509
; GENERAL INFORMATION:
; APPLICANT: SERRE, Guy Bruno Rene
; APPLICANT: SIMON, Michel
; APPLICANT: WEBER-VIVAT, Marina
; TITLE OF INVENTION: POLYPEPTIDE EXPRESSED IN THE HORNY LAYER OF EPIDERMIS
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 016800-336
; CURRENT APPLICATION NUMBER: US/09/381,656
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: FR 97/03899
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: FR 97/11317
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0 .
; SEQ ID NO 1
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-09-381-656-1

Query Match 100.0%; Score 149; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 5.1e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
||||||||||||||||||||||||||||||
Db 501 DILAQVKPLGPQLADPEVFLPQGELLDSP 529

RESULT 2
US-10-164-230-61
; Sequence 61, Application US/10164230
; Patent No. 6933118
; GENERAL INFORMATION:
; APPLICANT: Inoko, Hidetoshi
; APPLICANT: Tamiya, Gen
; TITLE OF INVENTION: METHOD OF TESTING FOR PSORIASIS VULGARIS
; FILE REFERENCE: 06501-112US1
; CURRENT APPLICATION NUMBER: US/10/164,230
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: PCT/JP00/08624
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: JP 11/346867
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 63

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 143
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 151
;   OTHER INFORMATION: Xaa = Ser or Cys
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 152
;   OTHER INFORMATION: Xaa = Ser or Arg
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 202
;   OTHER INFORMATION: Xaa = Ser or Phe
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 282
;   OTHER INFORMATION: Xaa = Pro or Thr
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 401
;   OTHER INFORMATION: Xaa = Ser or Gly
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 408
;   OTHER INFORMATION: Xaa = Ala or Ser
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 409
;   OTHER INFORMATION: Xaa = Gly or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 410
;   OTHER INFORMATION: Xaa = Ser or Leu
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 453
;   OTHER INFORMATION: Xaa = Ser or Asn
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 527
;   OTHER INFORMATION: Xaa = Asp or Asn
US-10-164-230-61

  Query Match             95.3%;  Score 142;  DB 2;  Length 529;
  Best Local Similarity   96.6%;  Pred. No. 5.7e-13;
  Matches   28;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLPQGELLDSP 29
              ||||||||||||||||||||||||||| ||
Db        501 DILAQVKPLGPQLADPEVFLPQGELLXSP 529


RESULT 3
US-09-755-456-10
; Sequence 10, Application US/09755456
; Patent No. 6890536
; GENERAL INFORMATION:
;  APPLICANT: DELBAC, FREDERIC
;  APPLICANT:  DANCHIN, ANTOINE
;  APPLICANT:  VIVARES, CHRISTIAN
;  TITLE OF INVENTION: MICROSPORIDIAN POLAR TUBE PROTEINS, NUCLEIC ACIDS
;  TITLE OF INVENTION:  CODING FOR THESE PROTEINS AND THEIR APPLICATIONS
;  FILE REFERENCE: 1566-00
;  CURRENT APPLICATION NUMBER: US/09/755,456
;  CURRENT FILING DATE:  2001-01-05
;  PRIOR APPLICATION NUMBER: PCT/FR99/01630
;  PRIOR FILING DATE: 1999-07-06
;  PRIOR APPLICATION NUMBER: FR 98/08692
;  PRIOR FILING DATE: 1998-07-07
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Encephalitozoon hellem
US-09-755-456-10

  Query Match             36.9%;  Score 55;  DB 2;  Length 272;
  Best Local Similarity   30.0%;  Pred. No. 2.7;
  Matches   15;  Conservative    3;  Mismatches    8;  Indels   24;  Gaps    1;

Qy          4 AQVKPLGPQLADPEVFLP----------------------QGELLDSP 29
              ||| |: || | |  |||                      ||:: | |
Db         14 AQVAPVTPQAAVPTQFLPGAQQKIGGVDNRCANKQVEGVQIFQGDMADCP 63


RESULT 4
US-09-949-016-8863
; Sequence 8863, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8863
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8863

  Query Match             35.2%;  Score 52.5;  DB 2;  Length 272;
  Best Local Similarity   57.9%;  Pred. No. 6.5;
  Matches   11;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy         10 GPQLADPEVFL-PQGELLD 27
              || :||| | | ||||:::
Db         38 GPPVADPGVALSPQGEIIE 56


RESULT 5
US-08-486-049-3
; Sequence 3, Application US/08486049
; Patent No. 6572862
;  GENERAL INFORMATION:
;    APPLICANT:  Estes, Mary K
;    APPLICANT:  Jiang, Xi
;    APPLICANT:  Graham, David Y
;    TITLE OF INVENTION:  Methods and Reagents to Detect and
;    TITLE OF INVENTION:  Characterize No. 6572862walk and Related Viruses
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  801 Pennsylvania Ave., N.W.
;      CITY:  Washington, D.C.
;      STATE:
;      COUNTRY:  USA
;      ZIP:  20004
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,049
;      FILING DATE:  June 7, 1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Davis, Peter
;      REGISTRATION NUMBER:  36,119
;      REFERENCE/DOCKET NUMBER:  311.023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-662-0200
;      TELEFAX:  202-662-4643
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  530 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-486-049-3

  Query Match             32.9%;  Score 49;  DB 2;  Length 530;
  Best Local Similarity   44.0%;  Pred. No. 47;
  Matches   11;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          5 QVKPLGPQLADPEVFLPQGELLDSP 29
              || |: | : :  |  ||||   ||
Db         48 QVNPIDPWIINNFVQAPQGEFTISP 72


RESULT 6
US-10-314-739A-3
; Sequence 3, Application US/10314739A
; Patent No. 6942865
; GENERAL INFORMATION:
;  APPLICANT: Estes, Mary
;  APPLICANT:  Jiang, Xi
;  APPLICANT:  Graham, David
;  TITLE OF INVENTION: Methods and Reagents to Detect and Characterize No. 6942865walk a
;  FILE REFERENCE: HO-P00486US7
;  CURRENT APPLICATION NUMBER: US/10/314,739A
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 08/486,049
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 07/696,454
;  PRIOR FILING DATE: 1991-05-06
;  PRIOR APPLICATION NUMBER: US 07/433,492
;  PRIOR FILING DATE: 1989-11-08
;  PRIOR APPLICATION NUMBER: US 07/515,933
;  PRIOR FILING DATE: 1990-04-27
;  PRIOR APPLICATION NUMBER: US 07/573,509
;  PRIOR FILING DATE: 1990-08-27
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 530
;   TYPE: PRT
;   ORGANISM: No. 6942865walk Virus
US-10-314-739A-3

  Query Match             32.9%;  Score 49;  DB 2;  Length 530;
  Best Local Similarity   44.0%;  Pred. No. 47;
  Matches   11;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          5 QVKPLGPQLADPEVFLPQGELLDSP 29
              || |: | : :  |  ||||   ||
Db         48 QVNPIDPWIINNFVQAPQGEFTISP 72


RESULT 7
US-08-686-528A-2
; Sequence 2, Application US/08686528A
; Patent No. 6054134
; GENERAL INFORMATION:
;   APPLICANT:  LINGWOOD, Clifford A.
;   TITLE OF INVENTION:  HAEMOPHILUS ADHESIN PROTEIN
;   NUMBER OF SEQUENCES:  4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Burns, Doane, Swecker & Mathis, L.L.P.
;     STREET:  1737 King Street, Suite 500
;     CITY:  Alexandria
;     STATE:  Virginia
;     COUNTRY:  United States
;     ZIP:  22314-2756
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/686,528A
;     FILING DATE:  26-JUL-1996
;     CLASSIFICATION:  424
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Rea, Teresa Stanek
;     REGISTRATION NUMBER:  30,427
;     REFERENCE/DOCKET NUMBER:  032609-001
;  INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  337 amino acids
;     TYPE:  amino acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-686-528A-2

  Query Match             32.6%;  Score 48.5;  DB 2;  Length 337;
  Best Local Similarity   43.3%;  Pred. No. 33;
  Matches   13;  Conservative    4;  Mismatches    6;  Indels    7;  Gaps    2;

Qy          1 DILAQVKPLG---PQLAD----PEVFLPQG 23
              |:|| |||||     :||     :| :| |
Db         24 DVLASVKPLGFIVSSIADGVTGTQVLVPAG 53


RESULT 8
US-09-456-287-2
; Sequence 2, Application US/09456287
; Patent No. 6218147
; GENERAL INFORMATION:
;   APPLICANT:  LINGWOOD, Clifford A.
;   TITLE OF INVENTION:  HAEMOPHILUS ADHESIN PROTEIN
;   NUMBER OF SEQUENCES:  4
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Burns, Doane, Swecker & Mathis, L.L.P.
;     STREET:  1737 King Street, Suite 500
;     CITY:  Alexandria
;     STATE:  Virginia
;     COUNTRY:  United States
;     ZIP:  22314-2756
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/456,287
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/686,528
;     FILING DATE:
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa Stanek
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  032609-001
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  337 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-456-287-2

  Query Match             32.6%;  Score 48.5;  DB 2;  Length 337;
  Best Local Similarity   43.3%;  Pred. No. 33;
  Matches   13;  Conservative    4;  Mismatches    6;  Indels    7;  Gaps    2;

Qy          1 DILAQVKPLG---PQLAD----PEVFLPQG 23
              |:|| |||||     :||     :| :| |
Db         24 DVLASVKPLGFIVSSIADGVTGTQVLVPAG 53


RESULT 9
US-09-652-603-2
; Sequence 2, Application US/09652603
; Patent No. 6753413
; GENERAL INFORMATION:
;  APPLICANT:  Ching, Yick-Pang
;  APPLICANT:  Wang, Jerry
;  TITLE OF INVENTION:  p35nck5a Binding Proteins
;  FILE REFERENCE:  N00/0295/US
;  CURRENT APPLICATION NUMBER: US/09/652,603
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  US 60/151,584
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 586
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-09-652-603-2

  Query Match             32.2%;  Score 48;  DB 2;  Length 586;
  Best Local Similarity   45.5%;  Pred. No. 75;
  Matches   10;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          5 QVKPLGPQLADPEVFLPQGELL 26
              | ||  |:: ||  :|   |||
Db         72 QEKPSSPEVEDPPPYLSGDELL 93


RESULT 10
US-09-758-759-79
; Sequence 79, Application US/09758759
; Patent No. 6861513
; GENERAL INFORMATION:
;  APPLICANT: Hosted, Thomas J.
;  APPLICANT:  Wang, Tim X.
;  APPLICANT:  Horan, Ann C.
;  TITLE OF INVENTION: Everninomicin Biosynthetic Genes
;  FILE REFERENCE: ID0983K US
;  CURRENT APPLICATION NUMBER: US/09/758,759
;  CURRENT FILING DATE:  2001-01-11
;  PRIOR APPLICATION NUMBER: US 60/175,751
;  PRIOR FILING DATE: 2000-01-12
;  NUMBER OF SEQ ID NOS: 204
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 692
;   TYPE: PRT
;   ORGANISM: Micromonospora carbonacea
;   FEATURE:
;   OTHER INFORMATION: evsA
US-09-758-759-79

  Query Match             32.2%;  Score 48;  DB 2;  Length 692;
  Best Local Similarity   47.6%;  Pred. No. 92;
  Matches   10;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DILAQVKPLGPQLADPEVFLP 21
              |  |:|   ||| |||   :|
Db        342 DPTAKVTDFGPQAADPTAKIP 362


RESULT 11
US-09-327-681-6
; Sequence 6, Application US/09327681
; Patent No. 6284495
; GENERAL INFORMATION:
;  APPLICANT:  Sato, Katsuaki
;  APPLICANT:  Usuda, Yoshihiro
;  TITLE OF INVENTION:  Method for producing nucleic acid substances
;  FILE REFERENCE:  OP853
;  CURRENT APPLICATION NUMBER: US/09/327,681
;  CURRENT FILING DATE:  1999-06-08
;  PRIOR APPLICATION NUMBER:  JP 10-165704
;  PRIOR FILING DATE:  1998-06-12
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 285
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-327-681-6

  Query Match             31.9%;  Score 47.5;  DB 2;  Length 285;
  Best Local Similarity   41.4%;  Pred. No. 38;
  Matches   12;  Conservative    4;  Mismatches    8;  Indels    5;  Gaps    1;

Qy          6 VKPLGPQLADPEVFLPQG-----ELLDSP 29
              |: ||  ||:||  || |     ::|  |
Db         84 VQTLGQSLANPERILPGGYVYLTDILGKP 112


RESULT 12
US-09-252-991A-27848
; Sequence 27848, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 27848
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27848

  Query Match             31.9%;  Score 47.5;  DB 2;  Length 425;
  Best Local Similarity   60.0%;  Pred. No. 61;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    7;  Gaps    2;

Qy          4 AQVKPL-GPQLADPEVFLPQGELLD 27
              ||| || |||       | ||||||
Db        200 AQVIPLVGPQA------LTQGELLD 218
```

```
RESULT 13
US-09-431-470-4
; Sequence 4, Application US/09431470
; Patent No. 6433249
; GENERAL INFORMATION:
;  APPLICANT: Duvick, Jon
;  APPLICANT:  Simmons, Carl R.
;  APPLICANT:  Crasta, Oswald R.
;  APPLICANT:  Folkerts, Otto
;  TITLE OF INVENTION: The Use of Beta-Glucosidase to Enhance
;  TITLE OF INVENTION:  Disease Resistance to Insects in Crop Plants.
;  FILE REFERENCE: 5718-43
;  CURRENT APPLICATION NUMBER: US/09/431,470
;  CURRENT FILING DATE:  1999-11-01
;  PRIOR APPLICATION NUMBER: 60/107,920
;  PRIOR FILING DATE: 1998-11-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 563
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-431-470-4

  Query Match             31.9%;  Score 47.5;  DB 2;  Length 563;
  Best Local Similarity   44.4%;  Pred. No. 85;
  Matches    8;  Conservative    6;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          7 KPLGPQLADPEVFL-PQG 23
              ||:|| : :| ::| |:|
Db        419 KPIGPPMGNPWIYLYPEG 436


RESULT 14
US-09-902-540-15577
; Sequence 15577, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15577
;   LENGTH: 418
;   TYPE: PRT
;   ORGANISM: Myxococcus xanthus
US-09-902-540-15577

  Query Match             31.5%;  Score 47;  DB 2;  Length 418;
  Best Local Similarity   45.5%;  Pred. No. 71;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          7 KPLGPQLADPEVFLPQGELLDS 28
              :| ||:|      :|||| :||
Db         58 QPRGPRL------IPQGEAMDS 73


RESULT 15
US-09-252-991A-23349
; Sequence 23349, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 23349
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23349

  Query Match             31.5%;  Score 47;  DB 2;  Length 427;
  Best Local Similarity   61.5%;  Pred. No. 73;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 DPEVFLPQGELLD 27
              |||: || |:|:|
Db        262 DPEIDLPAGQLVD 274


Search completed: March 21, 2006, 21:50:23
Job time : 26.8909 secs
```

This Page Blank (uspto)